Molecular epidemiology of Human Immunodeficiency Viruses and exploration of host immune responses associated with long-term control of HIV-2 infection in West Africa

Thesis submitted for the degree of Doctor of Philosophy

Thushan de Silva

University College London

Field site: Medical Research Council (UK) Laboratories, The Gambia

Supervisors:

Prof. Sarah Rowland-Jones, MA, DM, FRCP, FmedSci
Professor of Immunology
Nuffield Department of Medicine, Weatherall Institute of Molecular Medicine, Oxford, UK.

Prof. Robin Weiss, FRS
Professor of Viral Oncology
Division of Infection and Immunity, University College London, UK.

Dr. Matthew Cotten, PhD
Unit Virologist
Medical Research Council (UK) Laboratories, The Gambia.

I, Thushan de Silva, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGEMENTS

Without the constant support, encouragement and help of many individuals over the last four years, the challenge of undertaking a PhD across three countries, three laboratories and a rural West African field site would clearly have been an impossible task. My heartfelt thanks go to my supervisors, Sarah Rowland-Jones, Robin Weiss and Matt Cotten. Sarah, there are few clinical academics who reach the heights you have, yet still dedicate so much energy and time to enthusing medical students and junior doctors about science and supporting their early careers. Without your continued belief in the project and my ability, I would have given up a long time ago! Robin, thank you for welcoming me into your lab so readily and despite your busy schedule, taking the time to supervise yet another student. I feel fortunate to have had the opportunity to work with you and am grateful that you have maintained your interest in HIV-2 and Caió over all these years. Matt, your critical input into the work has been invaluable at all times, as well as your support when trying to work under challenging laboratory conditions. Thank you for having the patience to guide me through the last few years, having arrived in the Gambia, essentially as a clinician with limited laboratory skills. This work would also not have been possible without the generous support of a British Infection Society Research Fellowship and an MRC Clinical Research Training Fellowship.

During the process of setting up assays at the MRC Gambia, I have had to rely heavily on remote post-doctoral advice from many individuals in the UK. I am extremely grateful to Marlén Aasa-Chapman, Suzy Willey, Ed Wright, Anna Forsman and Willie Koh at UCL for always being willing to help during my visits to the Weiss laboratory, as well their prompt and useful email advice. Thank you to Tao Dong and Marie-Eve Blais at the WIMM in Oxford, for equally vital support and guidance during the T-cell components of my work. Many thanks also to others in Oxford who made my time there both productive and enjoyable: Louis-Marie, Jean, Peng, Karen, Sabelle, Nicky, Mohammed, Miguel, Godelieve, Julie, Eirini, Abi and Elham. I also acknowledge the generosity of James Robinson, Beatrice Hahn and Frederic Bibollet-Ruche in providing advice and vital laboratory reagents.

A big thank you to Stéphane Hué at UCL for taking the time and having the patience to educate me in phylogenetics and his willingness to respond to my never-ending email queries. I was lucky that you lack the ability to decline any request for help, despite clearly not having the time to do so!

My time at the MRC Gambia has been an unforgettable experience, made so by many friends and colleagues. Thank you to Clayton Onyango for his advice and help in the lab and constant positivity in the face of failing PCRs, to Irfan Zaidi for countless scientific discussions over coffee and help at any time of night and day with malfunctioning flow cytometers, to Martin Holland for sharing an office with me and putting up with being an easy target for ad hoc senior scientific advice, to Sarah Burr for inadvertently taking on this role during the last couple of years, to Katie Flanagan for her friendly encouragement and advice and providing inspiration, to Ramu Sarge-Njie for her excellent job running the HIV serology laboratory, to Abdoulie Jabang, Bakary Sanneh and Pa Saidou Chaw for their

efficient and high quality technical support, to Kevin Taylor, Shauna Mullaly and the entire biomedical engineering team who were often in the line of fire (and under appreciated) but worked tirelessly to keep the lab running under difficult conditions, to Jane Sutherland, Martin Ota, Bouke de Jong and Martin Antonio for graciously agreeing to share their Cat 3 space with me, to Safayet Hossin and Gilleh Thomas for excellent database support, to Assan Jaye for supporting the project, to Roxanne Turner and Roochi Trikha for extremely productive laboratory help during their placements in the Gambia, to Lucy Li and Harry Griffin for their help and excellent company during recruitment visits to Caió, and for showing remarkable ability beyond their experience in adapting quickly to running a field study and laboratory in rural conditions. Thank you also to all those at the Caió field station, who were always welcoming of visitors to the village and ever willing to do everything in their power to make a new study run smoothly: Tim Vincent for keeping the whole place running, as well as for hosting the best pizza and film nights in Guinea-Bissau, Dr. Mavinga for his clinical input, Moises and Djoaquim for their laboratory support, Natu for his fieldwork, Luis, Jordao and Djipan for help with study coordination as well as the ART programme, the project drivers Charles and Francoar for expertly navigating the sometimes flood ridden, hazardous terrain between the Gambia and Caió with precious reagents and samples in the back of their landrovers, and of course the Caió cohort members who still continue to participate in research projects after 20 years.

I am indebted to Hilton Whittle for his unshakable energy and enthusiasm in supporting all our work, and for knowing when to encourage and when to gently criticise to achieve the desired outcome. Many thanks to Peter Aaby for his willingness to always offer advice when asked. I am very grateful to Aleks Leligdowicz, without whom the Caió field station laboratory would not be the place it is now. Your incredible drive and achievements laid the foundations for so much further work. I also had the privilege of working with Carla van Tienen, who has contributed immeasurably to Caió and all the research based there over the last few years. Carla, without your constant input into the welfare of the field station and willingness to provide epidemiological support to all our Caió studies, I suspect the ART programme would have failed and I would certainly have struggled to complete my work. Thank you for your collaboration and friendship.

I had the pleasure of working in the GUM/HIV clinic at the MRC Gambia, under the supervision of Kevin Peterson. Kevin, thank you for your expert guidance on all clinical matters, as well as your thoughtful and detailed advice on everything else in life. Thank you also to Ken Joof and all the GUM fieldworkers, Francis Oko and Toyin Togun for being excellent colleagues.

Above all, thank you to my parents, who taught me at an early age, the value of hard work and not giving up despite initial failures; and just as importantly, to my wife Shami, who more often than not, put her own studies to one side to allow me to work, tolerated my difficult times, and always ensured family life kept functioning as it should. Finally, thank you to our little son Nayan, who allowed me to keep things in perspective when work was tough and made our time in West Africa even more enjoyable.

ABSTRACT

Unlike many parts of the world, both HIV-1 and HIV-2 are endemic in West Africa. The work in this thesis was based on a clinic cohort at the Medical Research Council Laboratories in the Gambia and a community cohort based in Caió, a rural village in Guinea-Bissau. Although HIV-2 was the dominant infection in these countries two decades ago, HIV-1 has increased rapidly in recent years. Chapter 3 demonstrates that although HIV-1 CRF02_AG is responsible for most HIV-1 infections in both cohorts, the subtype distribution varies considerably – including the presence of a novel circulating recombinant form (CRF49_cpx) in the Gambia, accounting for approximately 15% of all cases.

Chapter 4 applies modern phylogenetic and phylodynamic techniques to viral sequences from rural Guinea-Bissau to characterise the molecular epidemiology and population dynamics of HIV-1 and HIV-2 in this region. Viruses from HIV-2 elite controllers and progressors can share a most recent common ancestor, suggesting that host factors are responsible for the dichotomous outcomes observed in HIV-2 infection. Chapter 5 explores whether humoral immunity plays a role in HIV-2 non-progression. Neutralising antibody responses of remarkable magnitude and breadth are seen in the plasma of HIV-2 infected subjects, yet there is no direct relationship to viral control. Chapter 6 demonstrates that polyfunctional HIV-2 Gag-specific CD8+ T-cell responses are, however, associated with control of HIV-2. The phenotype of these CD8+ T-cells varies from what is thought to be important in HIV-1 control. In contrast to HIV-1, escape from both humoral and cellular immune responses may be limited in HIV-2.

The aim of this work was to help understand the spread and transmission of HIV-1 and HIV-2 in West Africa, as well as to provide insight into why most HIV-2 infected individuals behave as long term non-progressors, whereas others

progress to AIDS like their HIV-1 infected counterparts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIDS	Acquired Immune Deficiency Syndrome
ART	Antiretroviral therapy
AZT	Zidovudine
BF	Bayes Factor
BLAST	Basic Local Alignment Search Tool
bp	base pairs
cDNA	Complementary DNA
CMV	Cytomegalovirus
CRF	Circulating Recombinant Form
CSW	Commercial Sex Worker
CTL	Cytotoxic T-lymphocyte
DEAE	Diethylaminoethanol
DMEM	Dubelcco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
DRC	Democratic Republic of Congo
DTT	Dithiothreitol
EBV	Epstein-Barr Virus
EC	Elite Controller
ESS	Effective Sampling Size
FCS	Foetal Calf Serum
FEL	Fixed-effects Likelihood
FFU	Focus Forming Units
FMO	Fluorescence-minus-one
FSC	Forward Scatter
GALT	Gut-associated Lymphoid Tissue
GTR	General Time Reversible
GTR+I+G	GTR with proportion of invariable sites and substitution rate heterogeneity
GUM	Genito-Urinary Medicine
HIV	Human Immunodeficiency Virus
HIV-D	HIV-1/HIV-2 dual
HKY	Hasegawa-Kishino-Yano
HLA	Human Leucocyte Antigen
HTLV-1	Human T-Lymphotropic Virus-1
ICS	Intra-cellular cytokine staining
IFN-γ	Interferon-gamma
IL-2	Interleukin-2
IL-7	Interleukin-7
LAHDB	Los Alamos HIV Database
LB	Lysogeny Broth
LN2	Liquid Nitrogen
LPS	Lipopolysaccharide
LTNP	Long Term Non-Progressor

LTR	Long Terminal Repeat
Mab	Monoclonal Antibody
MCCT	Maximum Clade Credibility Tree
MCMC	Markov Chain Monte Carlo
MFI	Median Fluorescence Intensity
$MgCl_2$	Magnesium Chloride
MHC	Major Histocompatibility Complex
ML	Maximum Likelihood
MRC	Medical Research Council
MRCA	Most Recent Common Ancestor
mRNA	Messenger RNA
Nab	Neutralising Antibody
N_e	Effective Population Size
NNRTI	Non-nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleos(t)ide Reverse Transcriptase Inhibitor
nt	Nucleotides
PBMC	Peripheral Blood Mononuclear Cell
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PHA	Phyto-haemagglutinin
PI	Protease Inhibitor
PNLG	Potential N-linked Glycosylation
REL	Random effects Likelihood
RLU	Relative Light Units
RNA	Ribonucleic Acid
RPMI medium	Roswell Park Memorial Institute medium
RT	Reverse Transcriptase
SEB	Staphylococcal Enterotoxin B
SGA	Single Genome Amplification
SH	Shimodaira-Hasegawa
SIV	Simian Immunodeficiency Virus
SLAC	Single Likelihood Ancestor Counting
SSC	Side Scatter
TBE	Tris-Borate EDTA (Ethylenediaminetetraceticacid)
TCR	T-cell Receptor
tMRCA	Time of the most recent common ancestor
TNF-α	Tumour Necrosis Factor-alpha
TPHA	Treponema Pallidum Haemagglutination test
URF	Unique Recombinant Form
VL	Viral Load
VSV	Vesicular Stomatitis Virus
WT	Wild-type

CHAPTER 1: INTRODUCTION

Almost three decades have now passed since the discovery of HIV-1 as a cause of Acquired Immune Deficiency Syndrome (AIDS) [1] and it is estimated that over 60 million individuals have been infected globally to date [2]. Recent public health campaigns have helped reduce the number of new annual infections to 2.7 million in 2010, a fall from previous years and down 21% from a peak in 1997 [3]. Access to antiretroviral therapy (ART) has also expanded substantially, with an estimated 6.6 million individuals in low- and middle- income countries receiving HIV treatment at the end of 2010 [3]. Nevertheless, the number of HIV-infected individuals worldwide continues to rise (approx. 34 million by the end of 2010) and universal ART coverage has been achieved by only a handful of countries so far [3]. Furthermore, despite the dedication of considerable finances towards this pursuit, the goal of an HIV-1 vaccine remains elusive. A major obstacle to achieving this objective has been the lack of knowledge about host immune responses that might be capable of preventing or containing viral replication in natural HIV infection, in the absence of ART.

A closely related retrovirus, HIV-2, was isolated a few years later in 1986 from a patient from Cape Verde suffering from AIDS [4]. As progression to advanced disease occurs less frequently and HIV-2 has remained confined primarily to West Africa, far less is known about HIV-2 than HIV-1. Yet it is estimated that at least 1 million individuals are infected in West Africa [5] and ART is often suboptimal [6]. Studying HIV-2 has the potential to further the understanding of how a potentially pathogenic retrovirus can be effectively controlled by the host, perhaps through the induction of host immunity or by other modes of viral restriction.

1.1 The origins and diversity of HIV-1 and HIV-2

Several non-human primates species are natural reservoirs for Simian Immunodeficiency Viruses (SIVs), which are retroviruses within the *Lentivirus* genus that are closely related to the two HIVs [7-12]. As sequence data from these SIVs emerged, it became evident that both HIV-1 and HIV-2 have been introduced into humans via several independent cross-species transmission events from non-human primates [13] (Figure 1.1). There are currently four known separate lineages of HIV-1: groups M, N, O and P. Phylogenetic analyses provide good evidence to suggest that SIVcpz*Ptt*, which infects Central African chimpanzees (*Pan troglodytes troglodytes*) [14], gave rise to the main pandemic HIV-1 group M and the less prevalent group N (only 13 documented cases to date, all in Cameroon [15]). The close relatedness of HIV-1 group M and N sequences to SIVcpz*Ptt* isolates from Cameroon makes this the likely location of these zoonotic events [10, 16]. Interestingly, the recently discovered HIV-1 lineage, group P [17], most probably emerged from a gorilla origin (SIVgor), but the likely source of HIV-1 group O is unknown and could be from either SIVcpz or SIVgor. Group O is also rare, accounting for less than 1% of the global HIV-1 burden, with cases described primarily in Cameroon and Gabon [18]. It is not clear what events facilitated these cross-species transmissions, but hunting for the purpose of bushmeat could have provided the required exposure [19]. Molecular phylodynamic studies have dated the origin of the M and O groups to the beginning of the 20th century [20, 21]. Robust estimates of when groups N and P emerged are not yet available. One study has suggested that the origin of group P was between 1845 and 1989, with a more precise estimate not possible due to the current paucity of group P sequences [22].

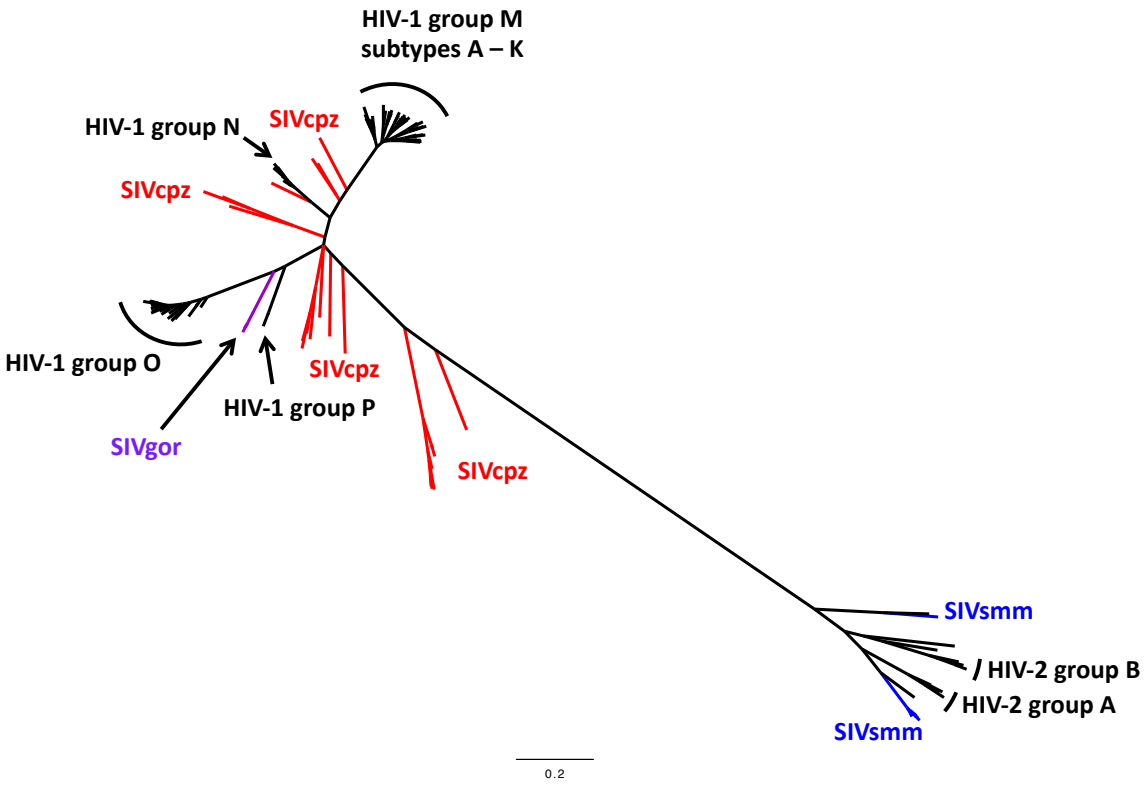


Figure 1.1. Bayesian phylogenetic tree showing the genetic relationship between SIVcpz, SIVgor, SIVsmm, HIV-1 and HIV-2 isolates based on polymerase gene sequences. Scale represents substitutions/site.

As the HIV-1 group M pandemic has evolved, the vast viral diversity that has been generated has come to light (Figure 1.2). Biological characteristics that contribute to this heterogeneity of HIV-1 include rapid viral turnover, an error-prone reverse transcriptase lacking proofreading ability and template switching of the reverse transcriptase between the two RNA genomes during reverse transcription [23-26]. HIV-1 group M has diversified into nine separate subtypes (A – D, F – H and K) and over 45 circulating recombinant forms (CRFs) [27]. The genomes of CRFs are mosaic and made up of fragments from different subtypes, such as CRF02_AG, the most prevalent HIV-1 subtype in West Africa (made up of subtypes A and G) [28]. When a recombinant form is made up of three or more pure subtypes, the suffix

'cpx' is used to denote a complex genome, such as CRF06_cpx [29]. As the distribution of different HIV-1 subtypes varies geographically, identification of local strains can help track the global HIV-1 epidemic. Over 50% of worldwide infections are due to infection with subtype C (the major form in Southern Africa and India), with subtype A (East Africa and Eastern Europe), subtype D (East Africa), subtype B (USA, Western Europe, Australia and South East Asia), CRF01_AE (South East Asia) and CRF02_AG (West Africa) responsible for a significant proportion of the rest [28, 30, 31]. It is likely that these differences in global subtype distribution largely reflect founder effects, where spread within a country, region or community takes place following the single introduction of an HIV-1 strain.

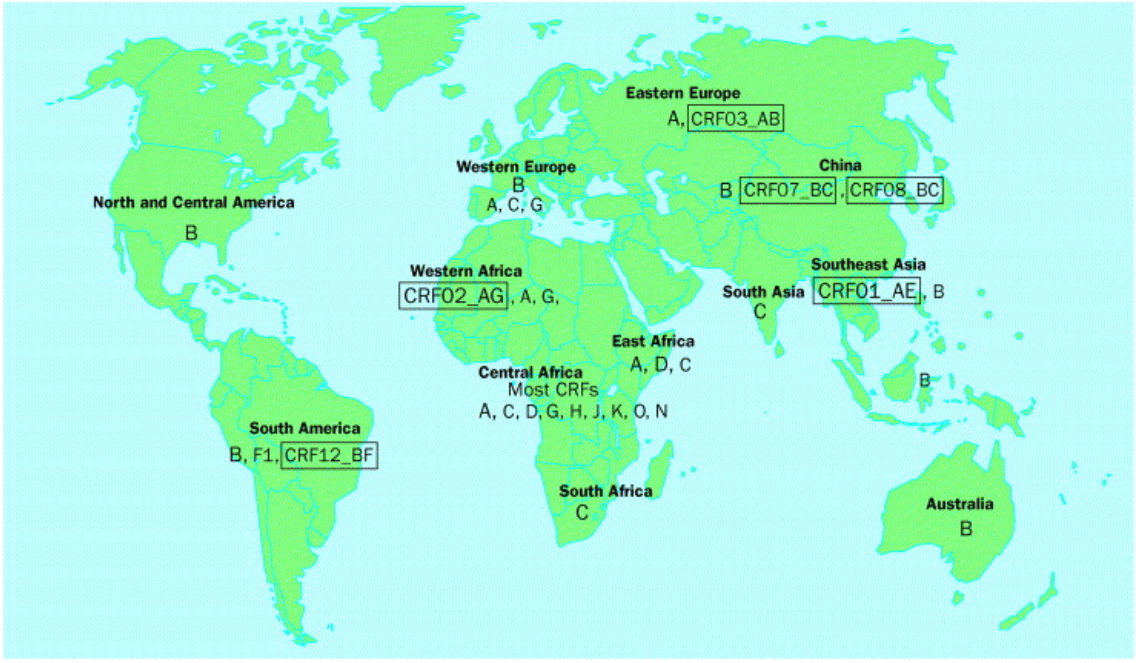


Figure 1.2. Global distribution of HIV-1 groups, subtypes and major circulating recombinant forms. Reproduced with permission from Perrin et al. Lancet Infect Dis 2003; 3(1): 22 – 27.

HIV-2 closely resembles SIVsmm from the West African sooty mangabey (*Cercocebus torquatus atys*) [7, 8, 32], a simian virus that is thought to have entered the human population on at least eight separate occasions, yielding eight distinct HIV-2 lineages A to H [33]. Of these, only groups A and B are endemic, with group A found throughout West Africa and group B prevalent mainly in Cote d'Ivoire [34-36]. Almost all the other groups are single-person transmission events that may have failed to spread beyond the initial human host: groups C, G and H are related to SIVsmm strains from Cote d'Ivoire, group D is similar to an SIVsmm from Liberia and groups E and F are most similar to SIVsmm strains from Sierra Leone [7, 32, 37, 38]. However, a second group F isolate has been recently discovered from a Sierra Leonean patient in New Jersey with features of HIV-2 disease progression [39]. Furthermore, an HIV-2 isolate from an 8-year-old boy from Sierra Leone was recently described, closely related to SIVsmm strains from the region but distinct from HIV-2 groups A to H [40]. Although this is yet to be formally classified as a novel HIV-2 group, it suggests that the knowledge of HIV-2 diversity and transmission in this region is incomplete. A recent report has also documented the first HIV-2 recombinant form (CRF01_AB), with a genome made up of group A and B fragments, that was isolated in Japan from two Africans (a Nigerian and a Ghanaian) and a Japanese individual [41]. Interestingly, a single isolate with the same recombination pattern was described in 1994 from Cote d'Ivoire (named 7312A), suggesting that CRF01_AB may be more widespread [42]. Although it is not known exactly how humans become infected with SIVsmm, a plausible route is again through the bushmeat trade. Sooty mangabeys are frequently hunted in West Africa and in one study, seven of 12 sooty mangabey bushmeat samples being sold in West African markets were infected with SIV [43]. The entry date of the HIV-2

subtype A ancestor into humans is estimated to be 1940 ±16 years, at least a decade after the introduction of HIV-1 group M [44].

1.2 The Epidemiology and transmission of HIV-2

Unlike the HIV-1 group M pandemic, HIV-2 has remained largely confined to West Africa, but has also spread to countries outside the sub-region via colonial links. While clinical cohorts exist in France and the United Kingdom, the largest number of cases outside West Africa are in Portugal [45]. HIV-2 has also been described in Mozambique, Angola, India and Brazil, which, along with Guinea-Bissau, all have ex-Portuguese colonial links [46-48]. Figure 1.3 shows African countries from where HIV-2 has been reported.

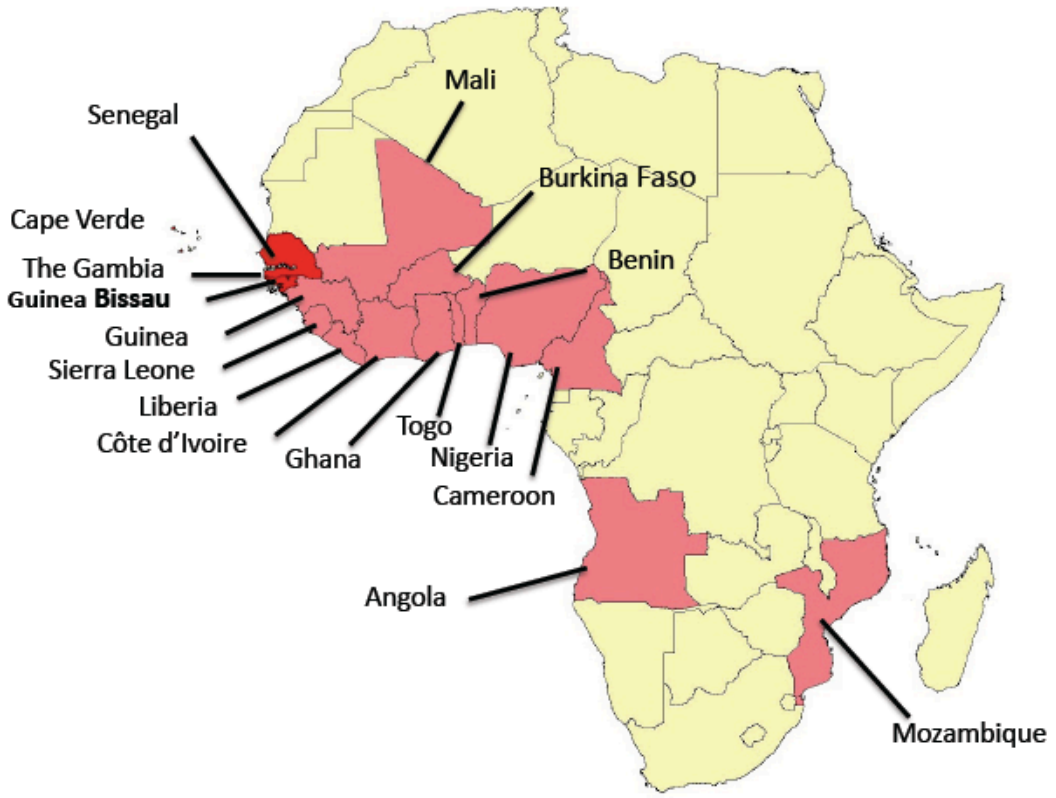


Figure 1.3. The epidemiology of HIV-2 in Africa. Countries with reported HIV-2 cases are coloured.

The highest prevalence of HIV-2 reported was two decades ago in Guinea Bissau: 8% in adults and up to 20% in individuals over 40 years of age [49]. Prevalence was highest amongst women in the 45-60 year age group, leading to speculation that older women are more susceptible to infection (possibly due to hormonal changes affecting the vaginal mucosa) [50]. Most countries (where reliable prevalence data are available) are now experiencing a decline in HIV-2 prevalence, whilst HIV-1 infection increases in the younger population [51-53]. For example, in a Senegalese sex worker cohort with a 20 year follow up, HIV-2 prevalence has fallen from 8% in 1985 to 5.5% in 2003, whilst HIV-1 increased from 1% to 13.8% in the same period [53]. The decline in HIV-2 prevalence is relatively straightforward to explain in view of its low transmission efficiency: for example, vertical transmission from HIV-2-infected breast-feeding mothers in the absence of ART is less than 4%, compared to up to 24.4% for HIV-1 in the same population [54]. Vertical transmission of HIV-2 is closely correlated with elevated plasma viral load (VL). Similarly, shedding of HIV-2 in genital secretions is significantly lower for HIV-2 compared to HIV-1 [55, 56] and the levels of HIV-2 in semen are closely related to plasma VL [56]. It is harder to understand why HIV-2 prevalence reached such heights in countries such as Guinea-Bissau in the 1980s in the first place. It has been postulated that increased prostitution, access to blood transfusions or other iatrogenic events (such as vaccination campaigns using unsterile injection equipment) during the protracted war of independence against Portugal could have facilitated transmission [57, 58]. The war might also have played a role in the spread of HIV-2 outside West Africa to Europe and to former Portuguese colonies such as Brazil and Angola. Even during this time of declining HIV-2 prevalence in West Africa, new cases still occur (most likely via heterosexual transmission) and not much is known regarding this ongoing spread of HIV-2.

Frequent asymptomatic infections may also mask the true prevalence of HIV-2 in the population. The incidence and prevalence rates from many West African countries are likely to be unreliable as many national HIV testing strategies do not incorporate the additional measures required to differentiate HIV-1, HIV-2 and HIV-1/2 dual infections.

1.3 The natural history and clinical features of HIV-2

Although an initial serological survey of asymptomatic Senegalese sex workers [59] raised the possibility that HIV-2 might be a non-pathogenic retrovirus, there were subsequent reports of clinical syndromes in HIV-2 patients indistinguishable from HIV-1 AIDS [60, 61]. The clinical features of HIV-2 AIDS mirror those in HIV-1, apart from the yet unexplained observation of less Kaposi's sarcoma with HIV-2 [62, 63]. In addition, presentation with AIDS and death tends to occur at higher CD4+ counts than for HIV-1 in the same clinical setting [62]. Evidence soon emerged, however, that the majority of HIV-2 infected individuals do not show signs of immunosuppression: following HIV-2 seroconversion, the AIDS-free survival at 5 years was significantly greater (100% vs. 67%) and the CD4+ T-cell decline was much slower than with HIV-1 [64-66].

Early studies from hospital cohorts in The Gambia confirmed greater survival with HIV-2 and two community cohort studies from Guinea-Bissau reported mortality risk in HIV-2 infected individuals as two-fold that of uninfected controls; lower than the mortality increase of over 10-fold described for HIV-1 infection in rural Tanzania [66-68]. A more recent study comparing both HIV-1 and HIV-2 associated mortality in rural Guinea-Bissau over two decades has confirmed these relative mortality risks [69]. HIV-2 associated mortality was less pronounced in older age

groups [66, 69], in stark contrast to the increasing risk of HIV-1 progression with age. This finding may be due to both a higher background mortality amongst older individuals, as well as a survival bias (cohort effect) of a group of older HIV-2 infected individuals who are long term non-progressors (LTNPs).

Further reports suggested the HIV-2 disease course might be heterogeneous, with some individuals progressing rapidly to AIDS and death. Although at high CD4 counts (>500 cells/μl or >28% of total lymphocytes) a clear survival advantage of patients with HIV-2 is evident over HIV-1 infected patients matched for disease stage, the mortality risk could be equivalent at intermediate and advanced immunosuppression [70, 71]. Development of assays capable of measuring HIV-2 plasma RNA load enabled further exploration of this apparent dichotomy. Whereas even HIV-1 LTNPs have a degree of detectable viraemia (other than the extremely rare 'elite controllers') [72], the majority of asymptomatic HIV-2 patients have undetectable plasma RNA [73, 74]. Many maintain this phenotype over almost two decades and show no greater risk of mortality when compared to age- and sex-matched HIV-uninfected controls [75]. In contrast, increasing baseline HIV-2 VL appears to predict greater risk of disease progression and death [71, 73, 75, 76]. The notion that HIV-2 infection uniformly progresses at a slower rate than HIV-1 infection is further refuted by the observation that when matched for baseline plasma VL, the median rate of $CD4^+$ T-cell decline per year is equivalent in both infections [77].

In contrast to HIV-1, the majority of HIV-2 natural history studies are based on cohorts where an exact date of infection is not known for most individuals. Primary HIV-2 infection has only rarely been described [78], although a study

capturing recently-infected individuals demonstrates that VL set points are lower in HIV-2 when compared to HIV-1 [79]. As mother to child transmission is also not common in HIV-2 [54], little is known about disease progression when HIV-2 is acquired vertically. However, the limited data available suggest that unlike most vertically HIV-1 infected children, HIV-2 infected infants may survive for many years (including into adulthood) without the requirement for ART [80, 81].

The proportion of individuals who have HIV-2 VLs below the limit of detection varies depending on the cohort studied. In the Caió community cohort in rural Guinea-Bissau, 37% of all subjects recruited at the inception of the cohort (in 1989) had undetectable VLs (<100 copies/ml) [75]. Follow up data from the French ANRS CO5 clinical HIV-2 cohort presents a contrasting picture. Using a definition of a LTNP as an HIV-2 infected individual infected for at least 8 years and a CD4 count of >500 cells/μl, only 6.1% of the cohort were LTNPs. Individuals with an HIV-2 VL below 500 copies/ml for at least 10 years made up only 9.1% of the cohort (all but one of these individuals had VL <100 copies/ml). While the obvious differences between a community and clinic-based cohort (where the latter will inevitably select for more symptomatic patients) may explain this discrepancy to some extent, it would be important to establish whether different host genetic or viral factors between cohorts also play a role.

Thus the study of HIV-2 raises an intriguing question: why does a human retrovirus with many similarities to HIV-1 result in asymptomatic infection and/or undetectable plasma VLs in most individuals, yet cause AIDS and death in others in a similar manner to HIV-1? Defective virological factors, better host immunity and host genetics are all possible explanations. The answer to this question could

provide important insight into HIV-1 pathogenesis and enhance the understanding of protective immunity to HIV infection.

1.4 Virological aspects of HIV-2 pathogenesis: is HIV-2 an attenuated virus?

Despite the significant difference in natural history of infected individuals, HIV-1 and HIV-2 share considerable sequence identity at the amino acid level (approximately 60% of Gag and Pol; 40% of Env) [82]. Unsurprisingly, the sequence homology between SIVsmm and HIV-2 is far greater (75 – 90%) [8]. Key differences in genome organization between HIV-1 and HIV-2 are the absence of a *vpu* accessory gene and additional presence of a *vpx* accessory gene in HIV-2 (Figure 1.4). It is now known that HIV-1 Vpu functions to counteract the human interferon-induced antiviral restriction factor tetherin by inducing its degradation [83]. In the absence of a *vpu* gene, HIV-2 (and SIVsmm) envelope proteins have acquired the ability to antagonize tetherin, albeit via a different mechanism (sequestration in the trans-golgi network) [84]. Recent studies have also characterised the role of HIV-2 Vpx in counteracting SAMHD1, a restriction factor operating in cells of the myeloid lineage, by inducing its ubiquitin-proteasome-dependent degradation [85, 86]. The lack of *vpx* in the HIV-1 genome results in restricted HIV-1 replication in myeloid cells, including dendritic cells. What relevance these genome structure differences have to the contrasting courses of HIV-1 and HIV-2 infection is not yet clear. In addition, an earlier study suggested that (primary myeloid and plasmacytoid) dendritic cells were less susceptible to infection by HIV-2 than HIV-1 [87]. It is difficult to reconcile these findings with the recently described role of Vpx, so further investigation is warranted.

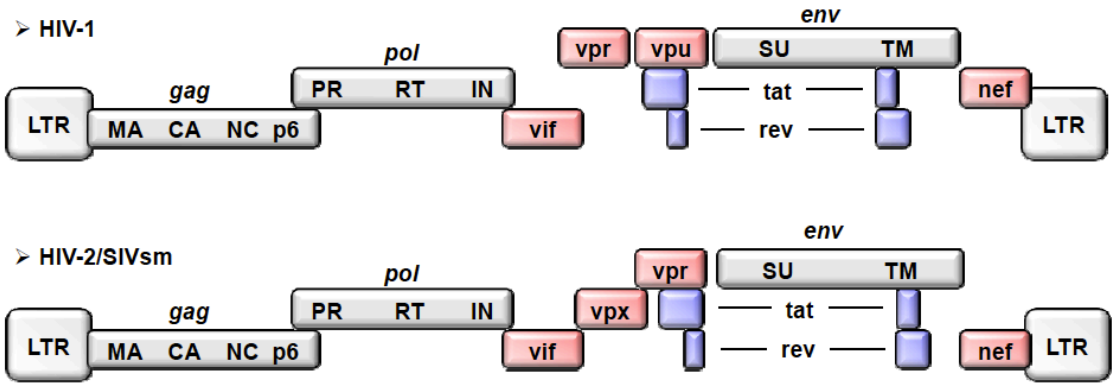


Figure 1.4. Genome structure of HIV-1 and HIV-2. Reproduced under the terms of the Creative Commons Attribution License [88] from Ayinde *et al.* Retrovirology 2010; 7:35. Structural genes are denoted in grey, regulatory genes in blue and accessory genes in pink. LTR = long terminal repeat, MA - matrix, CA - capsid, NC - nucleocapsid, PR - protease, RT - reverse transcriptase, IN - integrase. SU – surface protein and TM - transmembrane domain of *env*.

The slower progression of most HIV-2 infections could be explained by lower viral replication, enhanced host immune control, or both. Despite early suggestions that HIV-2 might be more attenuated, subsequent studies showed no difference in cytopathicity between HIV-1 and HIV-2 in a lymphoid histoculture model and suggested that coreceptor usage determines the cytopathic effect of both viruses [89, 90]. *In vitro* studies have also shown that most HIV-2 isolates are capable of using a broader range of coreceptors (e.g. CCR1, CCR4, GPR1) for cell entry than HIV-1, although it is not yet clear if the use of these alternative coreceptors is of significance *in vivo* [91, 92].

It is evident, however, that different *in vivo* viral dynamics prevail in the two infections. VL set points and plasma RNA levels during the asymptomatic stages of HIV-2 infection are approximately 30-fold lower than with HIV-1 [79, 93]. Proviral

loads are similar between HIV-1 and HIV-2 at corresponding levels of $CD4^+$ T-cell suppression [74, 94], but plasma VL is higher in HIV-1; except in those with advanced immunosuppression [79, 93-95]. What could explain these contrasting patterns of proviral and plasma VL in HIV-1 and HIV-2? Total cellular proviral DNA and integrated proviral DNA amounts are comparable, indicating that HIV-2 is equally able to establish a provirus [96]. Viral messenger RNA (mRNA) accumulation is lower in HIV-2 suggesting that the HIV-2 provirus has tighter transcriptional control [96]. HIV-2 could prefer to integrate in a reverse direction into transcriptional units and to integrate within heterochromatin (a rare event in HIV-1) [97], resulting in a higher likelihood of latent infection and lower viral transcription; although *in vitro* models might be a poor measure of the virus *in vivo* where complex factors are likely to influence integration and transcription. The HIV-2 long terminal repeat (LTR) is also different from HIV-1, with only a single NF-KB site, and could contribute to differential mRNA expression [82]. However, a recent study demonstrated similar levels of gag mRNA in HIV-1 and HIV-2 infected individuals matched for CD4 T-cell depletion [98], which implies that significant viral transcription occurs in HIV-2 regardless of the lower viraemia.

The lack of an HIV-2 pandemic might be explained by lower transmissibility secondary to lower circulating VLs, which *in vivo* is the net balance between replicative capacity (or fitness) and viral clearance. Is HIV-2 simply a less fit virus? HIV-2 isolates are less fit than HIV-1 group M viruses when compared by competition between viral strains in tissue culture [99]; although admittedly *ex vivo* fitness may not be a true reflection of fitness within human populations. The replication kinetics of isolates from aviraemic HIV-2 LTNPs are slower than HIV-2 progressor isolates and virus from HIV-1 LTNPs [100]. Similarly, a study of *env*

C2V3C3 sequences from asymptomatic HIV-2 patients over time shows low diversification and sequence divergence [101], demonstrating that HIV-2 turnover is usually extremely low, with the existence of a long-lived latent population. However, when envelope sequence data from HIV-2 and HIV-1 infected individuals with advanced disease (matched for CD4 T-cell count) were compared, HIV-2 *env* evolution was surprisingly faster than in HIV-1 [102].

The apparently dichotomous nature of HIV-2 infection outcomes has led to the suggestion that HIV-2 might represent a collection of related variants, of which only some are clearly pathogenic. As the majority of the HIV-2 viral sequence literature so far is inherently biased by isolates from HIV-2 infected individuals with higher VLs, this hypothesis has not been fully tested. The first study to explore HIV-2 viral sequence as a predictor of outcome examined HIV-2 proviral DNA sequences of *gag*, *pol, env* and LTR in the Caió community cohort in rural Guinea-Bissau [103]. A phylogenetic analysis using a joint alignment of LTR, *gag* and *pol* sequences suggested an association between viral genotype and survival, but not CD4 count or proviral load. This association was also not significant when alignments of the individual genes were considered. A more recent study, using plasma RNA-derived sequences from the same cohort, demonstrated a correlation between the presence of proline residues in three positions within the HIV-2 capsid and undetectable HIV-2 VL [104]. The mechanism behind this finding is not yet clear, but possibilities include greater susceptibility to the host restriction factor TRIM-5α [105], poorer capsid stability [104] and enhanced proteosomal processing (and therefore improved antigen presentation to capsid-specific T-cells) of HIV-2, in the presence of these proline residues (Sabelle Jallow, ms in preparation). These findings raise the question of whether infection with more or

less virulent strains can explain the different HIV-2 outcomes observed. Detailed phylogenetic studies are required to explore this possibility further. However, the finding that vertically transmitted HIV-2 (from mothers with high VLs and disease progression) can result in asymptomatic infection and long-term viral control in children [80] disprove this hypothesis.

1.5 Immune activation and HIV-2 pathogenesis

The observation that SIVsmm infection of the natural host, the sooty mangabey monkey, leads to high levels of plasma viraemia without evidence of immune suppression or disease progression has fuelled speculation that immune activation (which is characteristically low or absent in the SIV-infected sooty mangabey) rather than viral replication, is a fundamental mechanism by which HIV infection leads to disease. What is the relationship between immune activation and disease in HIV-2 infection?

The extent of immune activation in HIV-1 infection can be measured by the up-regulation of CD38 and HLA-DR on $CD4^+$ and $CD8^+$ T-cells or by monitoring soluble factors such as β_2-microglobulin. Chronic immune activation shows a significant correlation with disease prognosis in HIV-1 [106]. Similarly, in HIV-2 infection elevated β_2-microglobulin predicts death more reliably than VL [107]. The levels of systemic immune activation are similar in comparisons of HIV-1 and HIV-2 patients stratified by CD4 count, even though the plasma VL is significantly lower in the HIV-2 infected subjects with normal $CD4^+$ T-cell counts [108]. A recent study in the Caió community cohort demonstrated the direct correlation between immune activation markers and plasma VL in HIV-2, with most individuals with

undetectable viraemia displaying low levels similar to uninfected subjects [109]. Of note, however, were some subjects with undetectable VL with a significant degree of immune activation, suggesting mechanisms other than HIV-2 viraemia driven activation may exist.

What else might contribute to lower immune activation in HIV-2 infected subjects? One proposed mechanism is the ability of the SIVsmm/HIV-2 Nef protein to down-regulate the T-cell receptor (TCR)-CD3 complex on infected cells, which has been lost by HIV-1 [110]. This *nef* gene function could prevent over-activation of infected target cells and activation-induced cell death in the face of persistent virus replication [110]. It is possible, therefore, that variability in this Nef function could account for differences in immune activation (and therefore disease progression) in HIV-2 infected subjects. A study in the Caió community cohort found, as predicted, that the ability of Nef proteins from different donors to down-regulate CD3 correlated with immune activation, but there was no association seen with clinical outcome [111]. However, when only viraemic HIV-2 subjects are examined, a significant association is observed between Nef-mediated CD3 (and CD28) down-modulation and preserved CD4 counts [112]. It is possible, therefore, that Nef plays a role in limiting immune activation and disease progression in HIV-2 infected subjects in the face of persistent viraemia.

In HIV-1 infection, immune activation is thought to drive increased T-cell turnover whilst impaired thymic function means that the production of new T-cells cannot keep pace with T-cells lost through apoptosis. In contrast, T-cell apoptosis is reduced in HIV-2 infection relative to HIV-1 infection [113], most notably in asymptomatic subjects [114], which may relate to observations *in vitro* that HIV-2

isolates induce less apoptosis than do HIV-1 strains [115]. Reports suggest that thymic function appears to be enhanced in HIV-2 infection [116]: however, there is no elevation of circulating interleukin-7 (IL-7) levels in HIV-2-infected patients with a normal CD4 count, who also show normal levels of IL-7 receptor-α expression [117], suggesting better preservation of the ability to control T-cell homeostasis.

Acute infection with pathogenic SIV in rhesus macaques is characterised by massive infection of gut-associated lymphoid tissue (GALT) [118]. This might also provide a mechanism for immune activation as the damaged gut mucosa permits translocation of bacterial products such as lipopolysaccharide (LPS), which activates blood monocytes [119]. It will be interesting to explore the extent of GALT infection and lymphocyte depletion in HIV-2 infection and to determine whether increased gut permeability to microbial products or TCR downregulation is more important in immune activation and pathogenesis in HIV-2 progressors. It has recently been demonstrated that LPS concentration correlates with disease progression in both HIV-1 and HIV-2 infected individuals [120]. Intriguingly, acute SIVsmm infection of sooty mangabeys is also associated with profound GALT infection and T-cell depletion [121] but the levels of plasma LPS are low in chronic infection [119].

1.6 Host genetics and HIV-2 disease progression

As some HIV-1 infected individuals are LTNPs, host determinants must exist that influence the course of HIV-1 infection. The current evidence suggests that the key genetic determinants of HIV-1 control lie within the Major Histocompatibility Complex (MHC) and therefore affect class I peptide presentation [122]. These

variations can result in both improved control of HIV-1 viraemia or faster progression to AIDS. Little is known about the role of host genetic variability in susceptibility or disease progression in HIV-2. A small study in Senegal suggested that Human Leucocyte Antigen (HLA) B35 was associated with a greater risk of HIV-2 disease progression [123]. A more recent (and larger) study in the Caió community cohort failed to find any 'protective' HLA class I alleles, but demonstrated a significant association between HLA-B*1503 and HIV-2 progression with a low CD4 count and high VL [124]. Similar studies in other HIV-2 cohorts are necessary to explore these associations further.

1.7 Host immune responses in HIV-2

Neutralising antibody responses

Neutralising antibody (Nab) responses play a key role in inhibiting the biological activity of many viruses and therefore contribute to viral control. Unfortunately HIV-1 possesses the ability to replicate vigorously in the presence of a significant antibody response [125]. Despite the development of autologous Nabs in most HIV-1 infected subjects in early infection, HIV-1 gp120 is able to mutate considerably, allowing repeated escape from Nab responses whilst still retaining full function [126-129]. The resistance to neutralisation by primary HIV-1 isolates is explained in part by the structure of gp120 [130, 131]. The envelope is a trimeric structure made up of gp120-gp41 heterodimers, held together via conserved, but non-exposed surfaces on gp120 and is heavily glycosylated, thus protecting much of gp-120 from Nab. Key receptor-binding sites are masked by hyper-variable loops and not accessible to antibodies, or in the case of the coreceptor binding site, are exposed too briefly and in spatially difficult positions for antibodies to access. Most antibodies mounted against primary isolates in HIV-

1 infected individuals either lack breadth (i.e. the ability to neutralise a wide repertoire of viral variants) and potency, or are non-neutralising and may be directed against viral debris rather than infectious virus [132].

Significant advances have been made in the last two decades in the field of Nabs in HIV-1 infection. It is now clear that a greater proportion (approximately 20%) of HIV-1 infected subjects than previously thought develop a degree of breadth, albeit later in infection [133]. Several novel monoclonal antibodies (Mabs) have also been isolated from individuals with broadly neutralising sera, which show remarkable potency and breadth of neutralisation [134]. Determining the specificities of these Mabs, as well as characterising the structural interactions with HIV-1 Env are the first steps in new approaches to antibody-based HIV-1 vaccine development. Nevertheless, there is no convincing evidence that even when broad Nab responses are elicited in natural HIV-1 infection, for these responses to be associated with viral control and slower disease progression [135-138].

Broadly neutralising antibodies may be enhanced in HIV-2 infection when compared to HIV-1 infection. Although the presence of Nabs in HIV-2 was demonstrated in an early study, patient numbers were small and primarily lab-adapted isolates were used [139]. These findings have been replicated by a few studies using autologous and heterologous virus, although these studies are still limited by size, lack of clinical correlation and the use of non-standardized neutralisation assays [140-142]. HIV-2 might intrinsically have less ability to evade Nabs; perhaps due to fewer potential N-linked glycosylation sites compared to HIV-1, resulting in a more open HIV-2 envelope V3 domain [142]. Variability

between neutralisation sensitivity in HIV-2 isolates occurs [140] and could be related to the ability to enter cells independent of CD4 [143]. Increased neutralisation sensitivity of CD4-independent isolates indicates a more open coreceptor binding domain in HIV-2, although HIV-2 cross-neutralisation of HIV-1 still appears to require CD4 binding [144]. Interestingly soluble CD4 binding to HIV-2 CD4-dependent isolates increases neutralisation by Mabs mapping to a variety of Env epitopes, and not only the coreceptor binding site [143], suggesting globally different conformations of CD4-dependent and independent HIV-2 isolates; and implies that HIV might have evolved to use CD4 in part as a mechanism to shield conserved epitopes from host Nabs.

Few studies have explored the relationship between autologous neutralisation and progression in HIV-2, largely due to the technical difficulties in growing primary HIV-2 isolates. In a cross-sectional evaluation of nine HIV-2 patients, those with AIDS had lower neutralising titres than healthy patients [141]. The only longitudinal study of HIV-2 Nabs to date (using four subjects) showed little evidence of neutralisation escape over time and in contrast to HIV-1, patient sera from baseline were able to neutralise autologous virus isolated many years later [142]. The potential lack of neutralisation escape in HIV-2 is intriguing and could be related to the persistence of proviral sequences and lower replication rate of most HIV-2 isolates relative to HIV-1 [100, 101]. Recent work showing slow viral evolution in the HIV-2 envelope could also represent a lower pressure from Nabs in HIV-2 infection [101].

A more recent study concentrating on heterologous neutralisation in ART naïve patients suggested that HIV-2 induces Nabs greater in breadth but lower in

potency than HIV-1 [145]. The higher Nab breadth seen in HIV-2 might not be driven by viral diversification and neutralisation escape with time (as in HIV-1), but is perhaps due to greater conservation of HIV-2 envelope sequences between infected individuals. Alternatively, a broad host antibody repertoire in most HIV-2 infected individuals might be a consequence of a typically better preserved immune system. However, a recent study has revealed that irrespective of the lower viraemia, HIV-2 infected subjects (mostly with preserved CD4+ T-cells counts) have significant depletion of memory B-cells (which correlated directly with T-cell activation) to a similar degree to HIV-1 infected subjects [146]. The question of whether autologous Nabs in HIV-2 infection play a role in the slower disease progression and lower plasma VLs is unanswered and has the potential to inform rational HIV vaccine design immensely. The impact of non-neutralising antibodies with effector functions such as complement mediated lysis and antibody-dependent cellular cytotoxicity is even less explored in HIV-2 infection and might be a fruitful avenue of research.

T-cell responses

The role of CD8+ T-cells responses in controlling viraemia during acute HIV-1 infection is well established [147-151]. Several studies using rhesus macaques infected with SIV have also demonstrated a profound and rapid increase in viraemia (and associated disease progression) once CD8+ T-cells are depleted [152, 153]. The study of HIV-1 infected individuals who maintain VLs below the limit of detection ('elite controllers', ECs) can yield important insights into potential cellular immune correlates of protection. The enrichment of certain HLA class I alleles such as HLA-B*57 and HLA-B*27 in HIV-1 controller cohorts [122, 154, 155], as well as evidence for HLA-associated selection in HIV-1 sequences resulting

in viral escape [156-158], provide additional evidence for the importance of the cellular immune response in HIV control. The quality of the HIV-1 specific CD8+ T-cells may be key: while HIV-1 progressor CD8+ T-cells secrete interferon-γ (IFN-γ) alone in response to antigen stimulation, EC CD8+ T-cells are more likely to secrete multiple soluble factors (i.e. polyfunctional) and proliferate [159-162]. They are also able to rapidly upregulate perforin upon antigen-specific stimulation, as well as efficiently deliver granzyme B and kill HIV-infected CD4+ T-cells [163-165]. As ECs are rare and form <1% of most cohorts, it is challenging to perform such analyses and, importantly, as individuals are drawn from genetically heterogeneous populations, large numbers will be required to tease out significant protective features. A major complication is that not all ECs may share the same mechanism of control, thus further increasing the difficulty of identifying factors responsible for viral control.

Is the cellular immune response responsible for the exceptional HIV-2 control seen in many subjects? Early studies compared HIV-1 and HIV-2 infected individuals matched for disease stage and found no significant differences in the frequency, magnitude or phenotype of HIV-specific CD8+ T-cells [166-168]. Further work, however, suggested that HIV-2 infected subjects maintain both CD4+ and CD8+ Gag-specific T-cell responses that are more polyfunctional than in HIV-1 infection [169, 170]. As donors in these latter studies were highly selected for the ability to respond to Gag, it was not possible to conclude whether this greater polyfunctionality was associated with HIV-2 viral control. Few studies have been conducted comparing cellular responses in HIV-2 controllers and progressors. A small study demonstrated an association between more pronounced HIV-2 specific cytotoxic T-lymphocyte (CTL) activity and lower HIV-2 proviral load [171]. A

comprehensive analysis of cellular responses to the whole HIV-2 proteome, using IFN-γ ELISpots and overlapping peptides, in 64 HIV-2 infected donors in the Caió community cohort was more recently reported [172]. Significant differences were observed between the magnitude and the specificity of the response between HIV-2 ECs and viraemic subjects: individuals with undetectable VLs were more likely to target the Gag protein and the magnitude of these responses also showed an inverse correlation with plasma VL. Furthermore, Gag-specific responses were focused on a narrow region of the virus (the major homology region of the HIV-2 capsid, p26) and six immunodominant epitopes were identified [172]. As these studies were performed using whole peripheral blood mononuclear cells (PBMCs) and ELISpots, it was not evident if CD4+ and/or CD8+ T-cell responses were responsible for these observations.

Establishing what particular qualities of HIV-2 specific CD8+ T-cells are associated with viral control would also further the understanding of HIV pathogenesis. For example, the differentiation phenotype of virus-specific T-cells may be related to function and viral suppression. While some studies in HIV-1 infected subjects claim HIV-specific T-cells from viral controllers have a terminally differentiated phenotype (e.g. CD27-negative) [159], others argue this feature is associated with high VLs, greater immune activation and disease progression [173]. Initial data, focusing purely on CD8+ T-cells targeting an HLA-B*35 restricted HIV-2 capsid epitope, suggest that HIV-2 specific T-cells may be at an earlier stage of differentiation [174]. CTL clones derived from these individuals also showed remarkable functional avidity; that is, they displayed exceptional sensitivity to antigen and responded to extremely low peptide concentrations [174]. Although highly avid CD8+ T-cells are thought to play a role in controlling HIV-1 viraemia

[175, 176], in most HIV-1 infected subjects these T-cells do not persist beyond early infection [177]. Further detailed characterisation of HIV-2 specific CD8+ T-cells associated with viral control is warranted and would help in the development of immuno-therapeutic strategies aimed at inducing viral control in HIV-2 progressors.

1.8 Treatment of HIV-2 infection

Experience with antiretroviral treatment of HIV-2 infected patients is limited; this is partly due to the geographical restriction of HIV-2 patients to resource-poor settings where ART only recently became available. There are several potential difficulties encountered when treating HIV-2. Natural resistance of HIV-2 to non-nucleoside reverse transcriptase inhibitors (NNRTIs) used commonly to treat HIV-1 infection, possibly due to the presence of residues such as Y181L in wild-type (WT) HIV-2, has complicated the choice of ART regime for HIV-2 [178, 179]. All licensed ART drugs have been designed and evaluated for activity against HIV-1 and extrapolation to HIV-2 infected patients, as is current practice, is unlikely to be optimum. Although WT HIV-2 appears to be resistant to the only licensed entry inhibitor [180], reports suggest that integrase inhibitors show anti-HIV-2 activity; despite 40% heterogeneity between the HIV-1 and HIV-2 integrase proteins [181].

HIV-2 might be naturally less susceptible to zidovudine (AZT), although the relevance of this to the use of AZT containing ART regimes in HIV-2 is not clear [182]. HIV-2 might also have a different pathway to HIV-1 in acquisition of AZT resistance and have different patterns of mutations accumulating under drug pressure [183]; such as the more frequent selection for the multi-nucleoside reverse transcriptase inhibitor (NRTI)-resistance mutation Q151M [184]. HIV-2

displays reduced *in vitro* sensitivity to some protease inhibitors (PIs) such as amprenavir, atazanavir, tipranavir, indinavir and nelfinavir compared to HIV-1 [185]. For this reason, darunavir, saquinavir and lopinavir are the PIs of choice at present when treating HIV-2. Although baseline susceptibility of HIV-2 to lopinavir appears to be equivalent, selection of resistance mutations on therapy might occur more readily than in HIV-1 [186]. Natural polymorphisms in the HIV-2 protease gene exist which could facilitate the development of drug resistance and the number of mutations required to confer high level resistance to PIs appears to be considerably fewer for HIV-2 than HIV-1 [187]. In ART-naïve patients, a poorer response to ART could be seen in both initial VL decline and CD4+ T-cell reconstitution in HIV-2 patients than HIV-1 patients, with no difference in baseline CD4 count or proportion commenced on a PI-regime [188]. In a recent study in the Gambia, where a high proportion of HIV-2 infected subjects were treated with an optimal lopinavir-based regime, CD4 count reconstitution up to 24 months was also slower when compared to HIV-1 infected subjects (on an NNRTI-regime) [189]; although this did not effect viral suppression, morbidity or mortality. This slower CD4 recovery could potentially be secondary to ongoing viral replication below the limit of detection (and increased immune activation) – a phenomenon that has been recently demonstrated in HIV-2 patients despite treatment with ART [98].

It is clear that there is limited information on which to base accurate guidelines for HIV-2 treatment. Given the small size of most HIV-2 cohorts, further advances are unlikely without multi-centre controlled trials. If the same standard of evidence-based care is to be offered to HIV-2 infected patients, as is commonplace in HIV-1 treatment, these studies are urgently required. Furthermore, given the suboptimal

ART care available at present, novel HIV-2 immuno-therapeutic strategies aimed at inducing viral control could provide substantial clinical benefit.

1.9 Summary

HIV-2 infection presents an intriguing puzzle: why does a pathogenic retrovirus fail to cause disease in the majority of infected individuals, yet emulates HIV-1 in the ability to cause AIDS in others? Both virological and immunological aspects require further exploration to help answer this question, which may in turn provide insight into HIV pathogenesis and novel ways of preventing or controlling HIV-1. Although comparative studies of HIV-1 and HIV-2-infected individuals have yielded several features of host immunity that are selectively preserved, superior or distinct in natural HIV-2 infection, currently the most compelling evidence suggests a role for T-cell immunity in defining the HIV-2 controller phenotype. Detailed characterization of these potentially protective T-cell responses is required before anti-HIV-2 immuno-therapeutic strategies focused on enhancing T-cell immunity can be designed. Other factors such as broadly neutralising antibody responses, which are being fiercely researched for HIV-1 vaccine immunology, are yet to be confirmed in HIV-2 control. If preliminary reports are corroborated, identifying elements of the HIV-2 envelope-Nab interaction resulting in more potent neutralising responses could in turn aid the design of HIV-1 immunogens.

The data presented in this thesis was gathered firstly to establish whether humoral or T-cell immunity is paramount in naturally contained HIV-2 infection and to provide detailed characterisation of these host responses that may also give clues to the biology of HIV-2 *in vivo*. Furthermore, via generation of substantial HIV-2

env sequence data, new knowledge is gained into the properties of the HIV-2 envelope, as well as the ongoing transmission of HIV-2 in rural Guinea-Bissau and the relationship between infecting viral strains and likelihood of becoming an HIV-2 controller. Finally, as the number of HIV-2 infections wane in West Africa, HIV-1 is rapidly becoming dominant and also included are molecular epidemiological studies of HIV-1 that help understand the increasing HIV-1 epidemic in the Gambia and Guinea-Bissau.

CHAPTER 2: METHODS

2.1 Study cohorts

The studies described here are based on data and participants from two distinct West African HIV cohorts: a community cohort from Caió, Guinea Bissau and a clinical cohort based at the Medial Research Council (MRC) Laboratories the Gambia Genito-Urinary Medicine (GUM)/HIV clinic in Fajara, the Gambia (Figure 2.1). All work exploring the immunological correlates associated with HIV-2 viral control was undertaken using subjects and samples from the Caió community cohort. Caió is a rural village made up of 10 zones set amongst mangroves, cashew forests and rice fields in a geographically isolated setting. Subsistence farming is commonplace, but people also migrate to neighbouring towns and cities for work and education, as well as travel frequently to European countries such as Portugal and France. The population of the village is approximately 10,000 and over 95% belong to the Manjako ethnic group. The Manjakos form a genetically homogeneous tribe [124] and are distinct from their neighbours in their animistic belief system, with strong traditions surrounding earth and ancestral spirits influencing many aspects of daily life [190]. In 1989, a project jointly run by the MRC Laboratories, The Gambia and the Bandim Health Project, Bissau city, was established in Caió to study HIV-2 infection following the observation that a large number of HIV-2 infected Commercial Sex Workers (CSWs) attending a GUM clinic in Ziguinchor (Cassamance region, Senegal) were from Caió. Since the establishment of the cohort, three sero-surveys of the general adult population (≥ 15 years) have been conducted in 1989-1991, 1996-1997 and 2006-2007 [51], in which HIV status was determined. In 1991, a cohort of HIV-positive cases and

age/sex-matched HIV-negative controls was established following the initial sero-survey and additional subjects were recruited in subsequent sero-surveys. Further studies were conducted in 1996, 2003, 2006 and 2008 in which CD4+ T-cell subset analysis, VL quantification and clinical evaluations were performed on all available members of the case-control cohort, allowing long term follow-up and study of the natural history of HIV-2 infection [75]. Subjects have free access to medical care provided by the project's physician and ART for HIV became available in 2007 as part of the national AIDS program. The community prevalence of HIV-2 in Caió has decreased from 8.3% in 1990 to 4.7% in 2007, whereas the prevalence of HIV-1 has risen from 0.5% in 1990 to 3.6% in 2007 [51]. As one of the only community cohorts of HIV-2 infected subjects worldwide, Caió represents a unique cohort to base studies exploring immunological and virological correlates of protection from HIV-2 disease progression, less affected by the unavoidable selection bias usually present in clinical cohorts of HIV-infected individuals.

Aspects of the molecular epidemiology of HIV-1 and HIV-2 infection described were also performed on the Caió cohort, but additional HIV-1 subtyping was conducted on samples from the MRC Laboratories the Gambia GUM/HIV clinic. This clinic was located near the capital city Banjul, but drew patients from throughout the country. Confidential HIV testing, counseling and clinical care was offered at the clinic from 1987 (approximately 2000 HIV-infected subjects in active follow up by August 2009) and was also the Gambia's leading ART referral centre with around 400 patients on ART prior to cessation of clinic services by the MRC in 2010.

Ethical approval for the studies undertaken was obtained from the Gambian Government/MRC Joint Ethics Committee and from the Republic of Guinea-Bissau Ministry of Health.

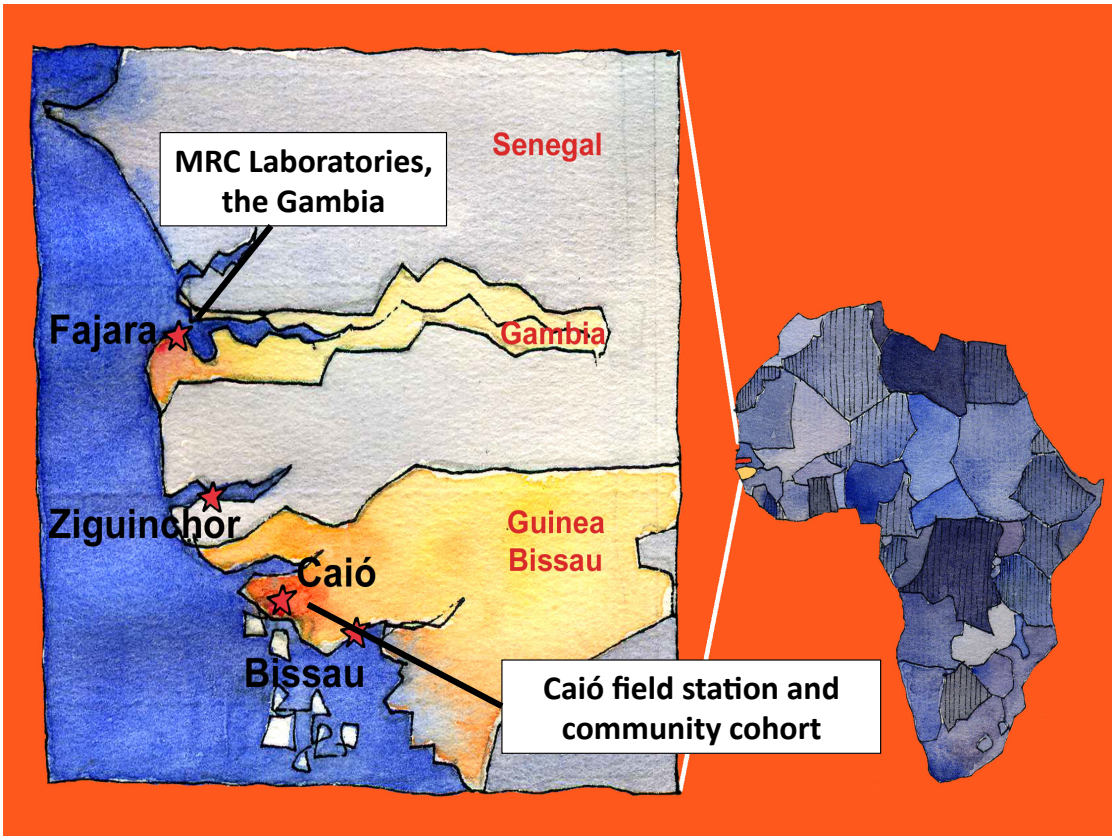


Figure 2.1. Map of West Africa illustrating location of MRC Laboratories the Gambia GUM/HIV clinic in Fajara and Caió community cohort in Guinea Bissau (by Matthew Cotten).

2.2 Extraction of Viral RNA

Viral RNA was extracted from 200µl of thawed plasma diluted in 800µl of RNase free water using the QIAamp Ultrasens Viral RNA Extraction Kit (Qiagen, UK) with final elution into 60µl. Each sample was loaded on a single column and washed according to the manufacturer's protocol. Briefly, 800µl of buffer AC was added to the 1ml of plasma/RNAse free water mix, along with 5.6µl of carrier RNA.

Following a 10 min incubation at room temperature (22°C - 25°C), the tubes were centrifuged at 1,200 x *g* for 3 min and the supernatant discarded. 300µl of buffer AR (prewarmed to 60°C) and 20µl of Proteinase K was added to the samples, followed by resuspension of the pellet and incubation at 40°C for 10 min. Buffer AB (300µl) was then added and the lysate applied to a QIAamp spin column for centrifugation at 4,000 x *g* for 1 min. The column was washed with buffer AW1 followed by centrifugation at 6,000 x *g* for 1 min and then buffer AW2 followed by centrifugation at 20,000 x *g* for 3 min. Finally, RNA was eluted in 60µl buffer AVE by centrifugation at 6,000 x *g* for 1 min. Extracted RNA was stored at -70°C until used in Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR) assays.

2.3 Amplification of HIV-2 and HIV-1 genes

Amplification of gp140 or gp160 HIV-2 env for neutralising antibody studies

HIV-2 *env* gene amplification was carried out in a nested RT-PCR using plasma-extracted viral RNA as template. Initial primers were designed by selecting conserved regions flanking the HIV-2 envelope from an alignment of all available HIV-2 isolates in the Los Alamos HIV Sequence Database and subsequent primer modification was done based on initial local sequence data obtained, which resulted in improved PCR efficiency. Final optimised primer sequences are showing in table 2.1. HIV-2 *env* products amplified were dependent on downstream methods of pseudovirus or replication-competent virus generation (see below): either full-length envelopes (gp160, approximately 2.5kb) or a fragment encoding the entire envelope ectodomain exposed to antibody recognition (gp140, approximately 2.1kb), but excluding the signal peptide and the gp41 cytoplasmic tail. Primers used in the nested PCR reaction for gp140

contained restriction recognition sites for *SnaBI* and *AgeI*, necessary for molecular cloning of these products into the 7312A-SNAG vector (see below).

Table 2.1 Primers used in HIV-2 *env* amplification and sequencing.

Primer	Function^	Sequence (5' to 3')
MO105	PCR *env* OF	GRGCMRGAGARCTCATTA
MO133	PCR *env* OR	CCTACTWGGTCATCATCATCTGAATCT
MO108	PCR *env* IF	CTYCTRCAYCAGACAAGTGAG
MO134	PCR *env* IR	GGTCATCATCATCTGAATCTACATCAT
MO148	PCR *env* OF	GTTTTTTAAACAAGGGGCTCGGGATA
MO167	PCR *env* IF	CCAAATACGTAACTGTTTTYTATGG*
MO168	PCR *env* IR	GGGAAGAGAAAACCGGTCTATAGCC*
MO080	PCR *env* OF	CAGTCATCACAGAGTCATGTG
MO076	PCR *env* OR	TCCTTGTGGATAYGAYCTGT
MO077	PCR *env* IR	GGAAGAGAAAACAGGACTATAGCC
MO124	*env* sequencing reverse	CTAGTGCCATTAAAGCCAAACCA
MO72	*env* sequencing and PCR IF	TCATGTGAYAAGCACTATTGGGA
MO122A	*env* sequencing forward	GTGGACTAACTGCAGAGGAGAATT
MO122.5	*env* sequencing forward	CAACAGCTGTTAGACGTGGTCAA
MO125A	*env* sequencing reverse	AGAAAACCCAAGAACCCTAGCAC

^OF – Outer Forward, OR = Outer Reverse, IF = Inner Forward, IR = Inner Reverse.
* Underlined are *SnaBI* restriction site in MO167 and *AgeI* restriction site in MO168.

Reverse transcription and the first round PCR were performed in a single reaction. Each 25µl RT-PCR contained the following mix: 1x PCR buffer Titan One Tube system (Roche Applied Science, Germany), 2.5mM magnesium chloride (MgCl2), 400nM dNTP mix, 0.5µM of each primer (MO105 and MO133 for gp160; MO148 and MO134 for gp140), 0.208 U/µl RNase inhibitor, 1µl of the Titan One Tube

enzyme mix and 5µl of extracted RNA. Reverse transcription was performed at 45°C for 45 min, followed by 95°C for 3 min, 10 cycles of 94°C (15 sec), 52°C (30 sec), 68°C (3 min), followed by 30 cycles of 94°C (15 sec), 52°C (30 sec), 68°C (3 min) plus 5 sec time extension at 68°C after each round and a final extension of 7 min at 68°C. The nested PCR used 1µl of the first-round RT-PCR product in 50µl containing: 1 x buffer (with 1.5mM MgCl2 final concentration), 0.05U/µl Expand HiFi Plus polymerase (Roche), 400nM dNTP mix and 0.5µM of each primer (MO108 and MO134 for gp160; MO167 and MO168 for gp140). For gp160 amplification, the reaction was conducted at 95°C for 3 min, 40 cycles of 94°C (15 sec), 56°C (30 sec), 72°C (3 min), and a final extension of 7 min at 72°C. For gp140 amplification, the reaction was conducted at 95°C for 3 min, 40 cycles of 94°C (15 sec), 52°C (30 sec), 72°C (3 min) and a final extension of 7 min at 72°C.

The PCR products were resolved on a 1% agarose (Tris-Borate EDTA (ethylenediaminetetraacetic acid), TBE) gel and DNA was visualised by ethidium bromide staining (0.1mg per 100ml gel volume) following gel electrophoresis at 100 volts for 60 minutes. The 2.5kb or 2.1kb products purified using the MinElute Gel Extraction Kit (Qiagen) according to the manufacturer's protocol (see section 2.4).

Amplification of partial HIV-2 env for phylogenetic and selection pressure analyses

Partial C2 – gp41 *env* genes (corresponding to HIV-2_{ROD}, accession number M15390, positions 6784 – 7859) were amplified in a nested RT-PCR using plasma-extracted viral RNA as template. Amplification was carried out using first round primers MO080 and MO76, followed by nested primers MO72 and MO77 (Table 2.1) and nested RT-PCR techniques described above. Conditions were as follows:

reverse transcription was performed at 50°C for 30 min, followed by 95°C for 3 min, 40 cycles of 94°C (15 sec), 54°C (30 sec), 72°C (2 min) and a final extension at 72°C (7 min). The second round reaction was conducted at 95°C for 3 min, 40 cycles of 94°C (15 sec), 56°C (30 sec), 72°C (2 min) and a final extension of 7 min at 72°C. Visualisation and purification of PCR products were performed as described above.

Amplification of HIV-1 env for subtyping and selection pressure analyses

Full length HIV-1 *env* amplification was carried out using plasma-extracted viral RNA as template. Reverse transcription and the first round of a nested PCR reaction were performed in single reaction. Each 25µl RT-PCR reaction contained the following mix: 1 x PCR buffer Titan One Tube System (Roche), 2.5mm MgCl2, 400nM dNTP mix, 0.1µM of primers O_*env*f and O_*env*r (Table 2.2), 0.208U/µl RNase inhibitor, 1µl of the Titan One Tube enzyme mix and 5µl of extracted RNA. Reverse transcription proceed at 45ºC for 45 min. followed by 95ºC for 3 min, 10 cycles of 94°C (30 sec), 56°C (30 sec), 68°C (3 min), followed by 30 cycles of 94°C (30 sec), 56°C 30 sec), 68°C (3 min) plus 5 sec time extension at 68°C after each round and a final extension of 7 min at 68ºC. The nested PCR reactions used 1 µl of the first-round RT-PCR product in 50µl containing: 1 x Buffer (with 1.5mM $MgCl_2$ final concentration), 0.05U/µl Expand HiFi Plus polymerase (Roche), 400nM dNTP mix, 0.25µM of primers MO130 and MO147 (Table 2.2). Amplification was conducted at 95ºC for 3 min followed by 40 cycles of 94°C (15 sec), 56°C (30 sec), 72°C (3 min), and a final extension of 7 min at 72ºC.

Partial HIV-1 C2 – gp41 *env* genes (corresponding to HIV-$1_{HXB2,}$ accession number K03455, positions 6584 – 7490) were amplified using first round primers CA*env*OF

and CA*env*OR, followed by nested primers CA*env*IF and CA*env*IR (Table 2.2). Reverse transcription and the first round PCR were performed in a single reaction. Each 25µl RT-PCR contained the following mix: 1x PCR buffer Titan One Tube system (Roche), 2.5mM MgCl2, 400nM dNTP mix, 0.5µM of each first round primers, 0.208U/µl RNase inhibitor, 1µl of the Titan One Tube enzyme mix and 5µl of extracted RNA. Amplification was conducted at 95°C for 3 min, followed by 40 cycles of 94°C (15 sec), 54°C (30 sec), 72°C (2 min) and a final extension at 72°C (7 min). The nested PCR used 1µl of the first-round RT-PCR product in 50µl containing: 1 x buffer (with 1.5mM MgCl2 final concentration), 0.05U/µl Expand HiFi Plus polymerase (Roche), 400nM dNTP mix and 0.5µM of nested primers. The second round reaction was conducted at 95°C for 3 min, 40 cycles of 94°C (15 sec), 56°C (30 sec), 72°C (2 min) and a final extension of 7 min at 72°C. Visualisation and purification of PCR products were performed as described above.

Limiting dilution PCR and Single Genome Amplification full length HIV-1 env

All full length HIV-1 *env* fragments (for the purpose of HIV-1 subtyping) were initially amplified using bulk PCR conditions on undiluted template and sequencing was carried out as described below for the highly variable V1/V2 region, followed by the entire *env* fragment if no double peaks were observed. In those samples showing multiple peaks in the V1/V2 region, the cDNA was then amplified using two different dilution methods in order to improve resolution of sequences. Both methods involved diluting the cDNA and running a standard PCR. First, three-fold limiting dilution of a single cDNA sample (reverse transcribed using the Titan One Tube RT-PCR reaction mix, for 45 min at 45ºC) was carried out (from 1:3 to 1:243), followed by the standard first round and nested PCR conditions as described above. The highest dilution at which the *env* fragment

amplification was successful was chosen for sequencing. If the V1/V2 region still contained multiple sequences, single genome amplification was carried out with a modified protocol to that described in the literature [191]. Briefly, three-fold dilution of complementary DNA (cDNA) was carried out with nine replicates per dilution (starting at the highest dilution at which the single sample limiting dilution PCR was successful), followed by the standard first round and nested PCR conditions as described above. An amplified *env* from the dilution where only one or two replicates yielded a positive PCR reaction (i.e. <30% of replicates positive [191]) was selected for sequencing and purified as described below.

Amplification of HIV-1 gag p24

Reverse transcription and the first round of a nested PCR reaction were performed in a single reaction. Each 50µl RT-PCR reaction contained the following mix: 1 x PCR buffer Titan One Tube System (Roche), 2.5mm $MgCl_2$, 200 nM dNTP mix, 0.5µM of primers MO042 or MO024 (alternate outer forward) and MO044 (Table 2.2), 0.208U/µl RNase inhibitor, 1µl of the Titan One Tube enzyme mix and 10µl of extracted RNA. Reverse transcription proceed at 50ºC for 30 min, followed by 95ºC for 3 min, 40 cycles of 94°C (30 sec), 54°C (30 sec), 72°C (1 min) and a final extension of 7 min at 72ºC. The nested PCR reactions used 1µl of the first-round RT-PCR product in 50µl containing: 1 x Buffer (with 1.5mM MgCl2 final concentration), 0.05U/µl Expand HiFi Plus polymerase (Roche), 400nM dNTP mix, 0.5µM of primers MO043 and MO045 (Table 2.2). Amplification was conducted at 95ºC for 3 min followed by 40 cycles of 94°C (30 sec), 56°C (30 sec), 72°C (1 min), and a final extension of 7 min at 72ºC. Visualisation and purification of PCR products were performed as described above.

Table 2.2. Primers used in HIV-1 amplification and sequencing. All HIV-1 primers were designed based on alignments of West African HIV-1 sequences from the Los Alamos HIV Sequence Database to select conserved sequences flanking regions of interest.

Primer	Function*	Sequence (5' to 3')
O_*env*f	*env* PCR OF	TYTCCTATGGCAGGAAGAAGC
O_*env*r	*env* PCR OR	TAACCCWTCCAGTCCCCCCTTTT
MO130	*env* PCR IF	GAGCAGAAGACAGTGGCAATGA
MO147	*env* PCR IR	CATCCMACTATRCTRCTTTTTGACC
MO150	*env* sequencing reverse	ATTCCATGTGTACYTTGTACTG
MO151	*env* sequencing forward	CAATTCCCATACATTATTGTGC
MO152	*env* sequencing reverse	CACTTCTCCAATTGTCCRTCAT
MO153	*env* sequencing forward	GACAAGCAATGTATGCCCCTCC
MO154	*env* sequencing reverse	ACCAATTCCACAYACTTGCCCA
MO155	*env* sequencing forward	CTGGAACKCTAGTTGGAGTAAT
MO042	*gag* p24 PCR OF	TAGTATGGGCAAGCAGGGAG
MO024	*gag* p24 PCR OF	AACCCACTGCTTAAGCCTCA
MO044	gag p24 PCR OR	TGCCAAAGAGTGATTTGAGGG
MO043	*gag* p24 PCR IF	TGYGTRCATCAAARGATAGA
MO045	*gag* p24 PCR IR	CCCCTTGYTGGAAGGCCA
MO034	5’ LTR to *gag* p24 PCR OF	TGAGCCTGGGAGCTCTCTG
MO186	p24 to *env* PCR OF	TTAARTGTTTCAACTGTGGCAAAGAAGA
MO187	p24 to *env* PCR OR	CAAGCATGKGTAGCCCAGAYATTATG
MO188	p24 to *env* PCR IF	ATGTGGGAARGARGGACACCAAATGAA
MO189	p24 to *env* PCR IR	TCCACACAGGTACCCCATAATAGACT
MO191	5’ LTR to *gag* p24 PCR OR	AATGCTGWRAACATGGGTATTACTTCTG
MO192	5’ LTR to *gag* p24 PCR IR	TCTATTACTTTYACCCATGCATTTAAAGT
MO193	*env* to 3’ LTR PCR IF	CAGACCCTTATCCCAAACCCAAC
MO194	*env* to 3’ LTR PCR IR	CCCCCCTTTTCTTTTAAAAAGWRGC
AJB-1R	p24 to *env* sequencing reverse	TATGGATTTTCAGGYCCAATTYTTG
AJB-2F	p24 to *env* sequencing forward	GCCCAAARGTTAAACAATGGCCA
AJB-3R	p24 to *env* sequencing reverse	TTCTGTATRTCATTGACAGTCCAGCT
AJB-4F	p24 to *env* sequencing forward	ACACCAGAYAARAARCATCAGAAAG

AJB-5R	p24 to *env* sequencing reverse	GATTCCTAATGCATACTGTGAGTCTG
AJB-6F	p24 to *env* sequencing forward	CAGACTCACAGTATGCATTAGGAATC
AJB-7R	p24 to *env* sequencing reverse	ACTAATTTATCTACTTGTTCATTTCCGCC
AJB-8R	p24 to *env* sequencing reverse	ATGTCTAYTATTCTTTCCCCTGCACTG
AJB-9F	p24 to *env* sequencing forward	ATTCCCTACAATCCCCAAAGMCARG
AJB-10F	p24 to *env* sequencing forward	TGATTGTGTGGCARGTAGACAGGAT
AJB-11R	p24 to *env* sequencing reverse	TCCATTCTATGGAGACYCCMTGACC
AJB-12R	p24 to *env* sequencing reverse	TGCCATAGGARATGCCTAAGCCYTT
AJB-13F	p24 to *env* sequencing forward	AARGGCTTAGGCATYTCCTATGGCA
Carpools	*pol* PCR OF	TCARATCACTCTTTGGCARCGACC
CApolOR	*pol* PCR OR	ACTGTCCATTTRTCAGGATGGAGYTCAT
CApolIF	*pol* PCR IF	ACCAAAAATGATAGGRGGAATTGGAGG
CApolIR	*pol* PCR IR	AGGATGGAGYTCATAHCCCATCCAAA
MO207	*pol* PCR OF	AAAARGGGCTGTTGGAAATGTGG
MO208	*pol* PCR OR	GARAGRCAGGCTAATTTTTTAGGGA

* OF – Outer Forward, OR = Outer Reverse, IF = Inner Forward, IR = Inner Reverse.

Amplification of near full-length HIV-1 genomes

In addition to *env* and *gag* p24 fragments, near full-length HIV-1 genome sequence was obtained by amplifying three further fragments: (A) 5' LTR to *gag* p24, (B) *gag* p24 to *env* and (C) *env* to 3' LTR. For fragment (A), reverse transcription and the first round of a nested PCR reaction were performed in single reaction. Each 25µl RT-PCR reaction contained the following mix: 1 x PCR buffer Titan One Tube System (Roche), 2.5mM MgCl2, 400nM dNTP mix, 0.5µM of primers MO034 and MO191 (Table 2.2), 0.208U/µl RNase inhibitor, 1µl of the Titan One Tube enzyme mix and 5µl of extracted RNA. Reverse transcription proceed at 50ºC for 30 min, followed by 95ºC for 3 min, 40 cycles of 94°C (30 sec), 54°C (30 sec), 72°C (1 min) and a final extension of 7 min at 72ºC. The nested PCR reactions used 1µl of the first-round RT-PCR product in 50µl containing: 1 x Buffer (with 1.5mM MgCl2 final

concentration), 0.05U/µl Expand HiFi Plus polymerase (Roche), 400 nM dNTP mix, 0.5µM of primers MO024 and MO192 (Table 2.2). Amplification was conducted at 95ºC for 3 min followed by 40 cycles of 94°C (30 sec), 56°C (30 sec), 72°C (1 min), and a final extension of 7 min at 72ºC. Fragment (C) was amplified with a nested PCR on products obtained with primers O_*env*f and O_*env*r as described above. The nested PCR reactions and conditions were identical to those used above for fragment (A), but using primers MO193 and MO194 (Table 2.2). For fragment (B), reverse transcription was performed in a 20µl reaction containing 1x Qiagen Long-range RT buffer, 1mM mix of each dNTP, 1µM of primer MO187 (Table 2.2), 0.04U/µl RNase inhibitor, 1µl LongRange Reverse Transcriptase (Qiagen) and 10µl of extracted RNA. Reactions were incubated at 42°C for 90 minutes followed by 85°C for 5 minutes. Each 50µl first round PCR contained the following: 1x Expand Long Template (Roche) buffer 1 (with 1.75mM MgCl2 final concentration), 400nM dNTP mix, 0.3µM of primers MO186 and MO187 (Table 2.2), 0.75µl of the Expand Long Template enzyme mix and 5µl of cDNA template. PCR conditions were as follows: 94ºC for 2 min, 10 cycles of 94°C (10 sec), 56°C (30 sec), 68°C (4 min), followed by 30 cycles of 94°C (10 sec), 56°C (30 sec), 68°C (4 min) plus 20 sec time extension at 68°C after each round and a final extension of 7 min at 68ºC. The nested PCR used 1µl of the first round PCR product in 50µl containing 1 x Expand Long Template (Roche) buffer 1 (with 1.75mM MgCl2 final concentration), 400nM dNTP mix, 0.5µM of primers MO188 and MO189 (Table 2.2) and 0.75µl of the Expand Long Template enzyme mix. Amplification was conducted using the same conditions as described above for the first round PCR. Visualisation and purification of PCR products were performed as described above.

Amplification of HIV-1 pol

HIV-1 *pol* fragments were amplified by nested RT-PCR using plasma-extracted RNA as template. Three combinations of outer and nested primer sets were used to maximise the number of samples successfully amplified (Table 2.2): (A) MO186/CApolOR followed by MO188/CApolIR (B) MO207/CApolOR followed by MO208/CApolIR and (C) CApolOF/CApolOR followed by CApolIF/CApolIR. Strategies (A) and (B) amplified protease to partial reverse transcriptase fragments representing positions 2034 – 3257 and 2071 – 3257 of HIV-1_{HXB2} respectively. Strategy (C) amplified a partial reverse transcriptase fragment representing positions 2381 – 3257 of HIV-$1_{HXB2.}$ Reverse transcription and first round of a nested RT-PCR were performed in a single reaction. The reaction mixture of 25µl for the first round contained nuclease free water, 5 X buffer, 400nM dNTP mix, 0.5µM of each outer primer, 40U/µl RNase inhibitor, 1µl enzyme mix, and 5µl of extracted RNA. Reverse transcription was performed at 50°C for 30 minutes followed by 95°C for 15 minutes and first round PCR with the following conditions: 40 cycles at 94°C (30 sec), 54°C (30 sec) and 72°C (2 min), with a final extension at 72°C for 10 min. 1µl of the first round RT-PCR product was used for the nested PCR reaction in 50µl which contained nuclease free water, 10 X buffer, 400nM dNTP mix, 0.5µM of each nested primer, and 5U/µl hotstar taq DNA polymerase. Amplification was performed at 95°C for 5 minutes followed by 40 cycles of 94°C (30 sec), 56°C (30 sec) and 72°C (2 min), with a final extension at 72°C for 10 minutes. Visualisation and purification of PCR products were performed as described above.

2.4 Gel extraction of PCR products

Following gel electrophoresis, products were purified using the MinElute Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. Briefly, the PCR

product was excised from the agarose gel and 3x volumes of QG buffer added (e.g. 300µl QG buffer to every 100µg gel weight). Following incubation at 50°C for 10 min (or until the gel dissolved), the sample was applied to a QIAquick column and centrifuged at 10,000 x *g* for 1 min to bind DNA to the column. 500µl buffer QG was added to the column, followed by centrifugation at 10,000 x *g* for 1 min to optimise removal of agarose traces. The column was then washed by addition of 750µl buffer PE and centrifugation at 10,000 x *g* for 1 min. Finally, DNA was eluted into 30µl RNAse/DNAse free water by further centrifugation at 10,000 x *g* for 1 min. DNA was quantified using a Nanodrop spectrophotometer and stored at -20°C if not used immediately.

2.5 Molecular cloning of HIV-2 *env* products for neutralising antibody studies

Purified full-length HIV-2 gp160 fragments amplified using primers MO108 and MO134 were cloned into the pcDNA3.1/V5-His-TOPO® vector (Invitrogen, UK). Addition of a poly-adenine tail was first performed using the following mix: 1 x PCR buffer GoTaq™ (Promega, UK) with 1.5mM final MgCl2 concentration, 400nM dATP, 0.175U/µl GoTaq™ (Promega) DNA polymerase and 5µl of the *env* purified PCR product. The mix was incubated for 10 min at 72°C. 4µl of this reaction was then added to 1µl of the manufacturer's salt solution (0.2M NaCl and 0.01M MgCl2 final concentration) and 1µl of the pcDNA3.1/V5-His-TOPO® vector and incubated at room temperature for 30 min. 1µl of this ligation mix was then added to XL2-Blue MRF' Ultracompetent cells (Agilent Technologies, UK), incubated on ice for 30 min and transformed by heat shock at 42°C for 30 seconds and placed back on ice for 2 min. 250µl of SOC media (Invitrogen) was then added and the mix placed in a shaking incubator at 30°C for 90 minutes. 200µl of the mix was subsequently

plated on Lysogeny Broth (LB) agar plates supplemented with 50μg/ml ampicillin and incubated overnight at 30°C.

Purified primary HIV-2 *env* gp140 fragments were cloned in to a full-length HIV-2 provirus 7312A-SNAG (in pBluescript SK+), constructed from a parental HIV-2 plasmid pJK7312A [42, 144] and modified to contain unique restriction sites *SnaBI* and *AgeI* flanking gp140 (a kind gift from Frederic Bibollet-Ruche and Beatrice Hahn). Both primary gp140 *env* products and the 7312A-SNAG vector were first digested with *SnaBI* and *AgeI* as follows: DNA (3μg of 7312A-SNAG or 1μg of gp140 *env*) was added to a mix containing 1 x buffer 1 (New England Biolabs, MA) and *AgeI* enzyme (15U for 7312A-SNAG reaction and 5U for gp140 *env*) and incubated for 1 hour at 37°C. Buffer 4 (New England Biolabs) at a final 1x concentration and *SnaBI* enzyme (15U for 7312A-SNAG reaction and 5U for gp140 *env*) was then added and the mix incubated for a further 1 hour at 37°C. The products were then resolved on a 1% agarose (TBE) gel, DNA was visualised by ethidium bromide staining and either the primary gp140 *env* products or the linearized 7312A-SNAG *env*-deficient fragment purified using the MinElute Gel Extraction Kit (Qiagen) as described above. Restriction enzyme digestion with *AgeI* and *SnaBI* was repeated on the linearized 7312A-SNAG *env*-deficient fragment to minimise potentially uncut plasmid and reduce contamination with parental 7312A-SNAG in downstream experiments. This product was then incubated for 1 hour at 37°C with 1 x Antarctic Phosphatase buffer (New England Biolabs) and 5 units of Antarctic Phosphatase enzyme. The enzyme was subsequently heat inactivated at 65°C for 5 minutes and products purified using the MinElute Gel Extraction Kit (Qiagen) as above. Purified patient gp140 *env* and *env*-deficient 7312A-SNAG fragments were mixed at a molar ratio of 5.5:1 (approximately 100ng

DNA of each), along with 1 x ligation buffer (New England Biolabs) and 400 cohesive end units of T4 DNA ligase (1µl) in a 20µl reaction, and incubated at 16°C for 16 hours. 1µl of this ligation mix was then added to XL2-Blue MRF' Ultracompetent cells (Agilent Technologies) and transformed as described above.

Following overnight incubation at 30°C, up to 96 colonies were picked from each plate for identification of colonies with HIV-2 *env* inserts by colony-PCR (for both gp160 and gp140 cloning). Using a sterile tip, a trace of each individual colony was inoculated into a separate reaction containing the following: 1 x PCR buffer GoTaq™ (Promega) with 1.5mM final MgCl2 concentration, 400nM dNTP, 0.1/µl GoTaq™ (Promega) DNA polymerase and primers MO167/MO168 at 0.5µM final concentration each. The reaction was conducted at 95°C for 3 min, 30 cycles of 94°C (15 sec), 52°C (30 sec), 72°C (3 min) and a final extension of 7 min at 72°C. Products were resolved on a 1% agarose (TBE) gel and DNA was visualised by ethidium bromide staining. Colonies containing inserts (i.e. positive for a 2.1kb HIV-2 *env* fragment on colony PCR) were inoculated in 5ml of liquid LB media supplemented 50µg/ml ampicillin and incubated overnight in a shaking incubator at 30°C. Plasmid DNA was purified using the QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer's protocol (see below) and quantified using a Nanodrop spectrophotometer. The presence of an insert was confirmed by restriction enzyme digestion with *SnaBI* and *AgeI* as described above for gp140 containing 7312A-SNAG constructs and with *BstX1* for gp160 inserts in pcDNA3.1/V5-His-TOPO® vector (Invitrogen) as follows: a reaction containing 1x Buffer D (Promega, UK), 0.1mg/ml acetylated Bovine Serum Albumin final concentration, 5 units of *BstX1* enzyme and 5µg of plasmid DNA was incubated at 50°C for 1 hour. A *BstX1* recognition sequence is found both immediately

upstream and downstream of the TOPO® cloning site in pcDNA3.1/V5-His-TOPO®. Digested plasmids were resolved on a 1% agarose (TBE) gel, DNA was visualised by ethidium bromide staining and the *env* insert-containing plasmids confirmed by the presence of either a 2.1kb (for gp140) or 2.5kb (gp160) fragment. As TA cloning with the pcDNA3.1/V5-His-TOPO® vector does not allow directional cloning, all gp160-pcDNA3.1/V5-His-TOPO® plasmids were sequenced with a T7 primer to exclude *env* fragments inserted in the reverse direction. A T7 promoter/priming site is found immediately upstream of the multiple cloning site in pcDNA3.1/V5-His-TOPO®. All remaining successfully cloned *env* fragments were transfected in 293T cells detailed in 2.11 and screened for functionality as detailed in 2.11.

2.6 Purification of plasmid DNA

Plasmid DNA was purified using the QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer's protocol. Briefly, 5ml of overnight bacterial culture in liquid LB media were pelleted by centrifugation at 17,900 x *g* for 5 min. The pellet was suspended in 250µl buffer P1, followed by addition of 250µl buffer P2 and incubation for 5 min to allow alkaline lysis to take place. 350µl of neutralisation buffer N3 was then added and the sample mixed thoroughly by repeated inversion. The solution was then centrifuged at 17,900 x *g* for 10 min, the resulting supernatant applied to a QIAprep spin column and centrifuged at 17,900 x *g* for 1 min to allow plasmid DNA to bind the column. The column was then washed using 750µl buffer PE and centrifugation for 1 min at 17,900 x *g*, followed by another centrifugation step after flow-through was discarded to remove remaining traces of buffer PE. Plasmid DNA was eluted into 50µl buffer EB (10mM Tris-Cal, pH 8.5),

quantified using a Nanodrop spectrophotometer and stored at -20°C until further use.

2.7 Sequencing of HIV-1 and HIV-2 genes

All sequencing reactions were performed by Macrogen [192] using the BigDye Terminator v3.1 cycle sequencing kit and an ABI3730XL sequencer. A list of primers used in HIV-1 and HIV-2 sequencing are detailed in Tables 2.1 and 2.2. Briefly, primers used were the respective nested PCR primers, followed by additional overlapping sequencing primers designed to hybridize to conserved regions approximately 600 – 800 base pairs (bp) apart. All primers for PCR and sequencing were synthesised by Metabion (Metabion International AG, Lena-Christ-Str. 44/I, 82152 Martinsried, Germany).

2.8 Assembly of HIV-1 and HIV-2 sequences

For all samples, the sequencing chromatograms were carefully inspected for sites of ambiguous sequence. All reliable sequence data were assembled using the BioEdit Sequence Alignment Editor [193] and aligned using the Cap Contig Assembly program. For each assembled sequence, the open reading frame (ORF) was established using alignments with relevant regions of either HIV-2_{ROD} or HIV-1_{HXB2} and the ORF finder in the Sequence Manipulation Suite [194]. In areas where premature stop codons appeared, the sequence chromatograms were re-examined to determine if miscalled nucleotides in the region could account for the loss of the open reading frame. Such errors were manually corrected to give full reads of the respective sequence.

2.9 Mammalian cell line maintenance

JC53-bl (TZM-bl) HeLa cells and NP2 cells were passaged in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) heat-inactivated Foetal Calf Serum (FCS) (Invitrogen). The Human Embryonic Kidney 293T cell line was passaged in DMEM + GlutaMAX™ (Invitrogen, UK) supplemented with 10% (v/v) FCS. All cell lines were grown at 37°C in a humidified atmosphere with 5% CO_2. Adherent cells were passaged every 3 – 4 days with the addition of trypsin (0.25%)-EDTA (1mM) (Invitrogen) prior to splitting cell monolayers. Between 1/3rd and 1/5th of the trypsinised cells were used to seed a new flask, depending on cell requirement for experiments. When passaging the NP2 cell line expressing the CD4 receptor and/or chemokine coreceptors CCR5 or CXCR4, the medium (DMEM + 10% FCS) was also supplemented with 100µg/ml of G418 (selection for CD4 receptor expression) and 1µg/ml puromycin (selection for chemokine receptor expression) as appropriate.

2.10 Transfection of Human Embryonic Kidney 293T cells

Transfections of HIV-2 *env* clones were performed in 24-well plates, 6-well plates, T25 flasks or T75 flasks depending on the purpose of transfection and quantity of viral stocks required. A day before transfection, 293T cells were plated at a density of 1 x 10^5, 4 x 10^5, 1.5 x 10^6 or 4.5 x 10^6 in 24-well plates, 6-well plates, T25 flasks or T75 flasks respectively. The following day, media was replaced before transfection. Fugene-HD (Roche) transfection reagent (1.2µl, 6µl, 11µl or 32µl) was premixed with plasmid DNA (500ng, 2µg, 4µg or 12µg) in serum-free Optimem (Invitrogen) (25µl, 100µl, 400µl or 1.2ml) for 15 min at room temperature and then added to each well or flask (6 well plate, 24 well plate, T25 flask or T75 flask respectively). Where 2-plasmid transfections for pseudovirus production were performed, the amount of envelope: backbone plasmid DNA was

used at a ratio of 1:2. After overnight incubation at 37°C, the medium was replenished and virus harvested at 48 hours. Harvested virus was centrifuged at 250 x *g* for 15 min to remove cellular debris prior to snap freezing 500µl aliquots in liquid nitrogen (LN2) for future use.

2.11 Screening for functional HIV-2 *env* clones

HIV-2 *env* clone functionality was assessed on TZM-bl cells obtained from the NIH AIDS Research and Reference Reagent Program (Cat. No. 8129). This is a HeLa cell clone naturally expressing the CXCR4 chemokine receptor and engineered to additionally co-express CD4 and CCR5 receptors. TZM-bl cells contain integrated firefly luciferase and *Escherichia coli* β-galactosidase reporter genes under regulatory control of an HIV-1 LTR, which are induced in *trans* by the HIV Tat protein following infection [195]. Viral stocks were diluted in duplicate at 1:5 and 1:25 in DMEM + 10% FCS (100µl/well) in white 96-well flat bottom plates (Nunc, Denmark). TZM-bl cells were then added (1 x 10^4/well in 100µl volume) in DMEM + 10% FCS containing Diethylaminoethyl (DEAE)-Dextran (Sigma, UK) at a final concentration of 10µg/ml to enhance infection (therefore final virus dilutions 1:10 and 1:50). Control wells with TZM-bl cells only (no virus) were included in the assay. TZM-bl cells were also cultured in clear 96-well flat-bottom plates to assess cell growth and condition by light microscopy. Following a 48 hour incubation at 37°C, 100µl of assay medium was removed from each well and 100µl of Bright-Glo luciferase reagent (Promega) was added. Cells were allowed to lyse for 2 min and luminescence was measured using a luminometer (Promega). Luciferase activity is quantified as relative light units (RLU). *Env* clones producing virus that resulted in at least 50x higher RLU in the 1:10 diluted well when compared to the cell-only well were chosen for expansion of viral stocks and neutralisation experiments.

2.12 Passage in peripheral blood mononuclear cells of viral stocks grown in 293T cells

Whole blood from 3 healthy lab donors (60ml each) was collected in EDTA and 30ml of blood layered on to 15ml of Lymphoprep™ (Axis-Shield, Norway) in 50ml Leucosep® tubes (Greiner Bio-One, UK). These were centrifuged for 15 min at 800 x *g* with the brake off. The lymphocyte interphase was harvested and washed twice in sterile phosphate buffered saline (PBS) by centrifuging at 250 x *g* for 5 minutes. The cells were resuspended at a concentration of 10^6/ml in Roswell Park Memorial Institute (RPMI)-1640 medium (Invitrogen) with 20% (v/v) heat-inactivated FCS, 2mM L-glutamine (Sigma, UK), 1% PenStrep (100U/ml penicillin and 0.1mg/ml streptomycin final concentration) (Sigma) (R20) and supplemented with phyto-haemagglutinin (PHA) (Remel Europe, UK) at a final concentration of 2.5µg/ml. Following incubation at 37°C for 72 hours, the cells were washed once in RPMI and resuspended at 10^6/ml in R20 supplemented with 5U/ml recombinant interleukin-2 (IL-2) (Peprotech, UK).

Virus infections were undertaken following a further 1 – 3 days of growth at 37°C. Cells were then washed once in RPMI, resuspended in R20+IL-2 (growth medium) and the number of viable cells counted using the trypan blue exclusion test. Approximately 10 x 10^6 viable cells were resuspended in 1ml growth medium and incubated with 500µl – 1ml of viral stock (representing 0.005 – 0.05 multiplicity of infection) propagated in 293T cells for 2 hours at 37°C after which the cells were washed in RPMI and resuspended in 10ml growth medium in a T25 flask. Seven days after infection, virus was harvested following removal of cellular debris by centrifugation, and snap frozen in LN2.

2.13 Titration of viral stocks

Viral titres were assessed on TZM-bl cells in white 96-well flat-bottom plates (Nunc). Virus stock was diluted in DMEM+10% FCS at 1:5 starting concentration and 5-fold serial dilutions up to 1:48828125 in quadruplicate (100µl/well). TZM-bl cells were then added (1 x 10^4/well in 100µl volume) in DMEM + 10% FCS containing DEAE-Dextran (Sigma) as described above. (therefore final virus concentrations 1:10 and 5-fold serial dilutions). Replicate wells with TZM-bl cells only (cell controls) were included in the assay. TZM-bl cells were also cultured in clear 96-well flat-bottom plates to assess cell growth and condition by light microscopy. Following a 48 hour incubation at 37°C, 100µl of assay medium was removed from each well and 100µl of Bright-Glo luciferase reagent (Promega) was added. Cells were allowed to lyse for 2 min and luminescence was measured using a luminometer (Promega). Viral titres were calculated as the amount of Tissue Culture Infectious Doses required to infect 50% of cells inoculated, present per ml of viral stock ($TCID_{50}$/ml), using a $TCID_{50}$ macro [196] based on a modified Reed-Muench method [197]. Wells with RLU greater than 2.5x the negative (cell) control were considered positive for infection.

2.14 Neutralisation assays on TZM-bl cells

Plasma samples were heat-inactivated by incubation at 65°C for 90 minutes. Plasma and HIV-2 human Mabs were 3-fold serially diluted in DMEM+10% FCS (100µl/well) in white 96-well flat bottom plates (Nunc). 200 $TCID_{50}$ of virus was added to each well in a 50µl volume and incubated for 1 hour at 37°C. The usual starting dilution of plasma and Mabs incubated with virus was 1:20 and 10µg/ml respectively. TZM-bl cells were then added (1 x 10^4/well in 100µl volume) in

DMEM + 10% FCS containing diethylaminoethanol (DEAE)-Dextran (Sigma) as described above. Assay controls included replicate wells with no plasma or Mabs (virus only control) and wells with TZM-bl cells only (cell control). TZM-bl cells were also cultured in clear 96-well flat-bottom plates to assess cell growth and condition by light microscopy. Following incubation at 37°C for 48 hours, 150µl of assay medium was removed from each well and 100µl of Bright-Glo luciferase reagent (Promega) was added. Cells were allowed to lyse for 2 min and luminescence was measured using a luminometer (Promega). Neutralisation titres were calculated as the reciprocal plasma dilution or concentration of Mab causing 50% reduction in RLU compared to the virus control (IC_{50}), following subtraction of background RLU from cell control wells. In instances where the reciprocal IC_{50} was greater than the highest dilution of plasma or lowest concentration of Mab tested, the assay was repeated with a higher starting dilution of plasma/lower starting concentration of Mab. Non-specific inhibition was assessed by testing all viruses against HIV-negative plasma from the same cohort of patients and all HIV-positive plasma samples against the Vesicular Stomatitis Virus (VSV) envelope (pseudotyped with the pSG3Δenv backbone, which expresses the entire HIV genome expect for Env, obtained via the NIH AIDS Research and Reference Reagent Program, Cat. No. 11051).

2.15 Coreceptor tropism and CD4 dependence assays on NP2 cells

NP2/CD4, NP2/CD4/CCR5, NP2/CD4/CXCR4, NP2/CCR5 and NP2/CXCR4 cells were maintained as described above and used for the following assays. Coreceptor tropism and CD4 dependence assays were performed either by using virus produced by transfection of chimeric gp140-7312-SNAG full length HIV-2 molecular clones (and subsequent immunostaining of NP2 cells) or pseudovirus

produced by co-transfection of gp160 *env* clones in pcDNA3.1/V5-His-TOPO® and pNL4-3.Luc.R-E- (and subsequent measurement of RLU). pNL4-3.Luc.R-E- contains the full HIV-1 genome, with the firefly luciferase gene inserted into *nef*, along with two frameshifts that render the clone deficient in Env and Vpr (obtained via the NIH AIDS Research and Reference Reagent Program, Cat. No. 3418).

Infection with chimeric HIV-2 virus and immunostaining of NP2 cells

One day prior to infection, cells were seeded at 2 x 10^4 per well in 48-well flat-bottom plates (Corning Incorporated, NY) and incubated overnight at 37°C. Viral stocks were diluted at 1:5 and 10-fold serial dilutions in DMEM+10% FCS and 100µl incubated with NP2 cells in duplicate for 2 hours at 37°C. The cells were washed once with medium and incubated with 500µl of DMEM+10% FCS per well for 72 hours at 37°C. The cells were monitored for syncitia formation during this time.

Following incubation, the medium from each well was removed and cells incubated with 500µl of ice-cold methanol/acetone mix (1:1) for 10 min to fix the cells. The cells were then washed twice with PBS and incubated with 200µl of heat-inactivated HIV-2 polyclonal sera (pooled from 10 Caió HIV-2 infected donor samples not used for *env* amplification or neutralisation assays) diluted 1:4000 in PBS+1% (v/v) FCS for 1 hour at room temperature. Cells were subsequently washed twice with PBS and incubated with 200µl of goat anti-human IgG monoclonal antibody conjugated with β-galactosidase (Cambridge Biosciences, UK) diluted 1:400 in PBS+1% FCS, for 1 hour at room temperature. The cells were then washed twice with PBS and incubated with 0.5mg/ml β-galactosidase substrate (X-gal in PBS containing 3mM potassium ferrocyanide, 3mM potassium ferricyanide and 1mM MgCl2) overnight at 37°C. Cells infected with HIV-2 develop

a blue colouration and individual or grouped stained cells were counted under light microscopy as focus-forming units (FFU). Dilutions with approximately 50 – 100 FFU were chosen to allow accurate counting and a mean of the two replicate wells taken to calculate viral titres (in FFU/ml) according to the following formula: FFU/(dilution factor x volume used for infection).

Infection with HIV-2 Env-pseudotyped virus

Pseudovirus produced by co-transfection of gp160 *env* clones in pcDNA3.1/V5-His-TOPO® and pNL4-3.Luc.R-E- were diluted and plated in quadruplicate as detailed in section 2.13. NP2 cells were then added at 1 x 10^4 per well and incubated at 37°C for 48 hours. Luminescence measurement and calculation of viral titre in $TCID_{50}$/ml on NP2 cells expressing different combinations of CD4/CCR5/CXCR4 receptors was performed as described in section 2.13.

2.16 Peripheral Blood Mononuclear Cell isolation, cryopreservation and transport for T-cell immunology studies

13ml of blood was collected in Caió, Guinea Bissau, from each HIV-infected donor in heparinised 15ml falcon tubes and layered on to 15ml of Lymphoprep™ (Axis-Shield) in 50ml Leucosep® tubes (Greiner Bio-One). These were centrifuged for 15 min at 800 x *g* with the brake off. The plasma layer was removed, centrifuged at 250 x *g* for 5 min to remove cellular debris, and stored in 2ml aliquots on dry ice. The lymphocyte interphase was harvested and washed twice in sterile PBS by centrifuging at 250 x *g* for 5 min. Cells were counted and resuspended in freezing medium (FCS supplemented with 10% (v/v) dimethyl sulfoxide (DMSO)) at approximately 5 x 10^6/ml. Graded PBMC cryopreservation was undertaken utilising a 'Mr Frosty' freezing chamber (Thermo Fisher Scientific, UK) containing

100% isopropan-2-ol. Cryovials containing PBMCs were placed in the freezing chamber and placed on dry ice overnight, prior to transferring to LN2. Cryopreserved PBMCs (in LN2) and plasma (on dry ice) were transported every 7 – 10 days from the Caió field station to long-term LN2 and -70°C freezer storage respectively at the MRC Laboratories in Fajara, the Gambia.

2.17 Quantification of CD4 counts, HIV-2 viral load, HIV serology and DNA extraction

2ml of blood was collected in Caió, Guinea Bissau, from each HIV-infected donor into EDTA. CD4 T-cell subset analysis was performed on-site at the Caió field station using whole blood staining with two-colour immunofluorescence reagents (anti-CD45/anti-CD4) and a CyFlow SL flow cytometer (Partec, Münster, Germany). Full blood count analysis was conducted on-site using a Medonic Haematology Analyser (Boule Medical, Stockholm). Blood was used to screen for *Plasmodium falciparum* infection with a malaria rapid diagnostic kit (PfHRP2) and if positive, infection confirmed with thick and thin blood smears. The remaining EDTA-anticoagulated blood was centrifuged at 250 x *g* for 5 minutes and the plasma layer removed and stored as above.

HIV serology was performed by the MRC Laboratories, the Gambia HIV diagnostic laboratory via well established protocols: screening was performed using the Murex ICE HIV-1.2.0 capture enzyme immunoassay (Murex Diagnostics, UK), reactive sera confirmed using HIV-1 and HIV-2 immunochromatographic test kits (HEXAGON HIV, Human Gmbh) and a synthetic peptide-based strip kit if required (Pepti-Lav 1-2, Sanofi Diagnostics Pasteur, France). Samples with indeterminate results after serological testing were subjected to HIV-1 and HIV-2-specific

diagnostic PCR assays performed on cellular DNA, using nested primers based on LTR regions specific for each retrovirus [73, 74]. HIV-2 plasma viral loads were quantified in the MRC Gambia HIV Laboratories using an in-house RT-PCR assay using specific LTR primers and detection of chemiluminescence as previously described [74]. The lower limit of detection for the viral load assay was 100 copies/ml of viral RNA.

2.18 Antibodies used for flow cytometric assays

The following directly conjugated Mabs were used: (i) CD107a-PECy5 (clone eBioH4A3), CD4-APCefluor780 (clone RPA-T4), IFN-γ-PECy7 (clone 4S.B3), Tumour Necrosis Factor-alpha (TNF-α)-efluor450 (clone Mab11), Interleukin-2 (IL-2)-PE (clone MQ1-17H12), CD27-APCe780 (clone 0323), CD160-efluor647 (clone BY55), (Ebioscience); (ii) CD8-ECD (clone SFCI21Thy2D3), CD45RO-TexasRedPE (clone UCHL1) (Beckman Coulter); (iii) Perforin-FITC (clone D48) (Diaclone, Gen-Probe); (iv) CD3-Qdot655 (clone S4.1), CD4-Qdot605 (clone S3.5), CD8-Qdot705 (clone 3B5) (Invitrogen); (v) CD14-V500 (clone M5E2), CD19-V500 (clone HIB19), Granzyme B-Alexafluor700 (clone GB11), Ki67-Alexafluor488 (clone B56), Bcl-2-V450 (clone Bcl-2/100) (BD Biosciences); (vi) CD38-APC (clone HIT2), 2B4(CD244)-PECy5.5 (clone C1.7), PD-1-PECy7 (clone EH12.2H7), T-bet-Brilliant Violet (clone 4B10), HLA-DR-Alexafluor700 (clone L243) (Biolegend). LIVE/DEAD® fixable Aqua Dead Cell Stain (for 405nm excitation) was obtained from Invitrogen.

Unconjugated CD57 Mab (clone NK-1, BD Biosciences) was conjugated in-house to Quantum dot 565 obtained from Invitrogen according to the manufacturer's protocol. Briefly, the CD57 Mab was concentrated to 1mg/mL in 300µl PBS using a

50 KDa molecular weight cutoff centrifuge concentrator and 6.1µl dithiothreitol (DTT) added to reduce the antibody at room temperature for 30 min. 125µl of Quantum dot nanocrystals were activated by addition of 14µl SMCC solution at room temperature for 1 hour. Following equilibration of desalting columns with exchange buffer, both the reduced Mab and activated quantum dots were applied to two separate columns and eluted into 500µl each. The Mab and Quantum dots were incubated together at room temperature for 1 hour to allow conjugation to take place. A further 30 min incubation at room temperature was then undertaken following addition of 10µl of 10mM 2-mercaptoethanol to quench the conjugation reaction. The reaction was then concentrated to approximately 40µl volume by centrifugation at 4,000 x *g* in an ultrafiltration device, followed by application to a separation column to elute into PBS. The conjugated antibody was stored at 4°C prior to further use. All antibodies were pre-titred in an iterative process to obtain optimum volumes required, until full panels to be used in final experiments were complete.

2.19 Intra-cellular cytokine staining following peptide stimulation

Thawing cells

Cryopreserved PBMCs were thawed rapidly by drop-wise addition of 10ml of warmed FCS. Following centrifugation at 250 x *g* for 5 min, cells were incubated in RPMI supplemented with 10% (v/v) heat-inactivated FCS, 2mM L-glutamine (Sigma), 1% PenStrep (100U/ml penicillin and 0.1mg/ml streptomycin final concentration) (Sigma) (R10) and 60µg/ml DNase solution (Type IV, Sigma) for 15 minutes at 37°C. Cells were washed and resuspended in R10, and viable cells counted by trypan blue exclusion test. Cells were rested overnight at a

concentration of 10^6/ml, after which cells were washed once in R10, viable cells quantified as above and cells resuspended in R10 at 10^6 viable cells/ml. Only samples with cell viabilities ≥ 80% were used in intra-cellular cytokine staining (ICS) assays. The majority (60/62) samples had cell viabilities of 90 – 99% after resting overnight.

Overlapping Gag and Env synthetic peptides

Overlapping peptide sets representing the entire HIV-2 Gag and Env proteins (n = 70 for Gag and n = 107 for Env), as previously described [172], were used for peptide stimulation and ICS. 15 to19-mer peptides, overlapping by 10 amino acids, were used following modifications made to include only amino acids known to be C-terminal anchor residues for the F pocket of MHC-I. The original (designed) 18-mer peptide was either shortened by 3 or lengthened by 1 amino acid(s) if an unfavourable amino acid was present at the C-terminus, an accepted strategy for dealing with this problem [198]. Peptide sequences were based on an HIV-2 subtype A consensus sequence derived from four HIV-2 genomes available in the Los Alamos National Laboratory HIV database (D194, SBL/ISY, MCR35 and MCN13, accession numbers X52223, AR363717, AY509260 and AY509259 respectively).

Intra-cellular Cytokine Stimulation assay

Cells were aliquoted at 5 x 10^5 in 500µl R10 and stimulated with an overlapping Gag peptide pool at 2µg/ml (each peptide) and co-stimulatory antibodies anti-CD28 and anti-CD49 (BD Bioscience, UK) at 1 µg/ml each. Cells stimulated with *Staphylococcus* Enterotoxin B (SEB, Sigma) at 2µg/ml (positive control) and co-stimulatory antibodies alone (negative control) were included in the assay, as well as unstimulated cells for an unstained control and SEB-stimulated cells for flow-

minus-one staining for perforin to aid with gating during flow cytometric analysis. In samples where sufficient cells were available, cells stimulated with an overlapping Env peptide pool at 2μg/ml (each peptide) was also included. Following incubation for 1 hour at 37°C, protein transport inhibitors Brefeldin A and Monensin (Ebioscience, UK, final concentrations 3μg/ml and 2μM respectively) and pre-titred anti-CD107a-PECy5 were added to stimulations. Following a further 5 hours incubation at 37°C, cells were washed using FACS wash buffer (PBS supplemented with 2% (v/v) FCS, 1mM EDTA and 0.05% sodium azide), supernatant discarded after centrifugation, cells resuspended in the residual volume and incubated for 10 minutes in the dark at 4°C with LIVE/DEAD® fixable Aqua Dead Cell Stain (for 405nm excitation). Surface stains with pre-titred anti-CD4-APCefluor780 and anti-CD8-ECD were then added and cells incubated for a further 30 min at 4°C. Cells were then washed and fixed/permeabilized using Cytofix/Cytoperm™ (BD Biosciences). Briefly, 250μl Cytofix/Cytoperm™ was added and the sample incubated for 20 min at 4°C. Following permeabilization, cells were washed with Perm/Wash™ buffer (BD Biosciences) and stained intra-cellularly with pre-titred anti-IFN-γ, TNF-α, IL-2 and perforin antibodies for 30 min at room temperature. Anti-perforin staining was excluded in the perforin-minus-one condition. Following staining, cells were washed twice in Perm/Wash™ buffer, once in FACS wash buffer and resuspended in FACS wash buffer prior to analysis by flow cytometry.

2.20 Flow cytometry following intra-cellular cytokine staining assays

Cells were analysed using a CyAn flow cytometer (Dako, Beckman Coulter). Flow cytometric compensation for fluorescence 'spill-over' was performed using the BD

CompBead Anti-Mouse Ig, κ set (BD Biosciences) and antibodies used for cell staining described above. Briefly, antibodies were added to separate tubes containing 1 drop each of Anti-Mouse Ig beads and the negative control beads (which do *not* bind κ light chain-bearing immunoglobulins). Following a 30 min incubation at 4°C, beads were washed and resuspended in FACS buffer. Between 300,000 and 400,000 events were acquired in the lymphocyte gate for all samples. Data analysis was performed using FlowJo version 9.1 (Treestar, San Carlos, CA). Forward scatter (FSC)-area *versus* FSC-height parameters were used to exclude cell doublets, followed by exclusion of dead cells and subsequent gating on the lymphocyte population using FSC and side scatter (SSC) properties. CD4+ and CD8+ T-cells were then identified, with the CD8+ gate designed to include $CD4_{dim}CD8_{bright}$ cells, as this population has been shown previously in HIV-infected patients to represent an enriched anti-viral subpopulation of CD8+ T-cells [199]. For functional analysis, gates were created for each measured marker (CD107a, IFN-γ, TNF-α, IL-2 and perforin). Boolean/combinatorial gate arrays were then created using the FlowJo platform, which determined the frequency of each possible response pattern based on all possible combinations of the markers of interest (see figure 2.2 for example of gating strategy). Responses detected in the negative (anti-CD28/anti-CD49d) control were subtracted from responses observed in stimulated samples for each response pattern individually prior to further analysis. In cases where the value in the stimulated sample is smaller than that of the anti-CD28/anti-CD49d control, negative values arise following background subtraction. These were set to zero, but to avoid systematic bias with this approach, a threshold was set below which all positive values were also zeroed on the assumption that the distribution of background values is symmetric around zero [200]. This threshold was chosen based on the distribution of negative

values for each dataset (CD8+ and CD4+ responses) following background subtraction of twice the 90th centile. This value was 0.009 for the CD8+ dataset and 0.01 for the CD4+ dataset. In addition, only individuals with a total response to Gag stimulation of ≥ 0.1% (of CD8+ or CD4+ T-cells respectively) were included in the analysis of polyfunctional responses.

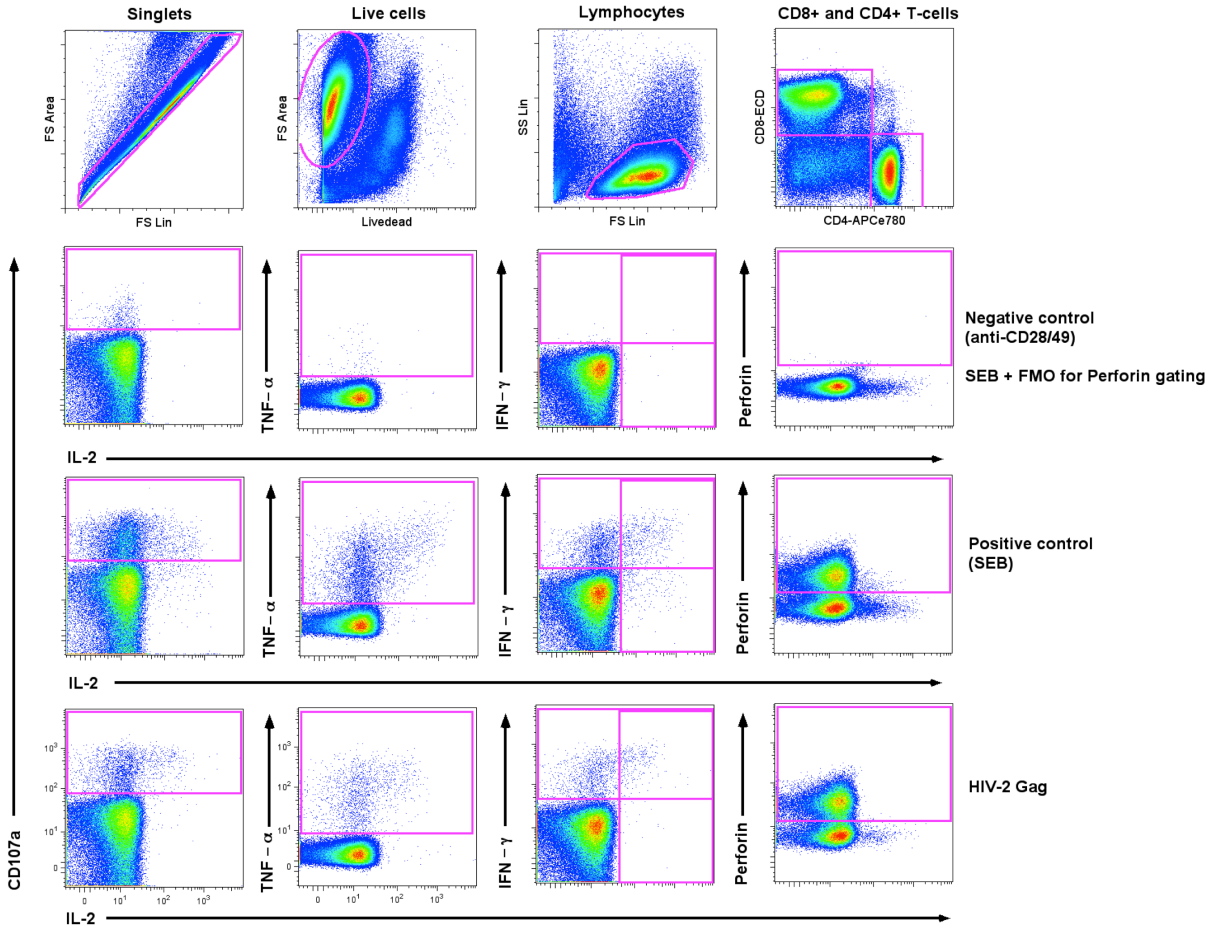


Figure 2.2. Example of flow cytometric gating strategy for intra-cellular cytokine staining following HIV-2 Gag stimulation. Shown are plots from one representative donor, with demonstration of gating for cytokine/soluble factor expression in CD8+ T-cells. FMO = fluorescence-minus-one.

2.21 Peptide-MHC class I tetramers

Monomeric complexes (50µg) of peptide/MHC class I, β2-microglobulin and biotin were freshly conjugated to streptavidin-PE (Sigma-Aldrich) to produce tetramer complexes. The tetramers used were as follows: HLA-B*3501-NPVPVGNIY, B*5801-TSTVDEQIQW, B*5301-TPYDINQML, B14-DRFYKSLRA (all HIV-2 Gag p26); A*0201-SLYNTVATL (HIV-1 Gag p17), A*0201-ILKEPVHGV (HIV-1 Pol), B8-FLKEKGGL (HIV-1 Nef), B8-DIYKRWII (HIV-1 Gag p24); A*0201-NLVPMVATV (CMV pp65), B35-EPLPQGQLTAY (EBV BZLF1) and B8-RAKFKQLL (EBV BZLF1).

2.22 Tetramer staining assays

Cryopreserved cells were thawed rapidly by drop-wise addition of warmed FCS. Following centrifugation at 250 x *g* for 5 min, cells were resuspended in R10 and 60µg/ml DNase solution (Type IV, Sigma-Aldrich) for 15 minutes at 37°C. Cells were washed and 1 x 10^6 live PBMCs (determined by trypan blue exclusion) used for each condition. Cells stained with Aqua for 30min at 4°C, washed and stained with 1µg of tetramer-PE for 15min at 37°C. Following 2 further washes, surface staining with CD3-Qdot655, CD4-Qdot605, CD8-Qdot705, CD14-V500, CD19-V500, CD27-efluor780, CD45RO-TexasRedPE, CD57-Qdot565 and PD-1-PECy7 were performed in all test samples. Additional surface staining with 2B4-PECy5.5 and CD160-efluor647 in one panel (i), and CD38-APC and HLA-DR-Alexafluor700 in another panel (ii) were performed. Following fixing and permeabilization as described above, intracellular staining with Perforin-FITC, Granzyme B-Alexafluor700 and T-bet-Brilliant Violet in panel (i) and Ki67-Alexafluor488 and Bcl-2-V450 in panel (ii) were done. FMO samples were included to aid gating during subsequent flow cytometric analysis.

Compensation for fluorescence 'spill-over' was performed using the BD CompBead Anti-Mouse Ig set (BD Biosciences) and antibodies as described above. Between 500,000 and 800,000 events were acquired for each sample. Data analysis was performed using FlowJo version 9.1 (Treestar). FSC area and FSC height were used to exclude cell aggregates, followed by removal of CD14+, CD19+, dead cells and gating on CD3+ cells. FSC and SSC were then used to gate on small lymphocytes, followed by isolation of CD8+ cells and tetramer-positive cells for further gating to examine expression of specific markers (Figure 2.3a and b). Similar approaches have been shown to reduce background staining during detection of tetramer-positive populations [201]. Boolean/combinatorial gate arrays were created to measure the co-expression of PD-1, 2B4 and CD160.

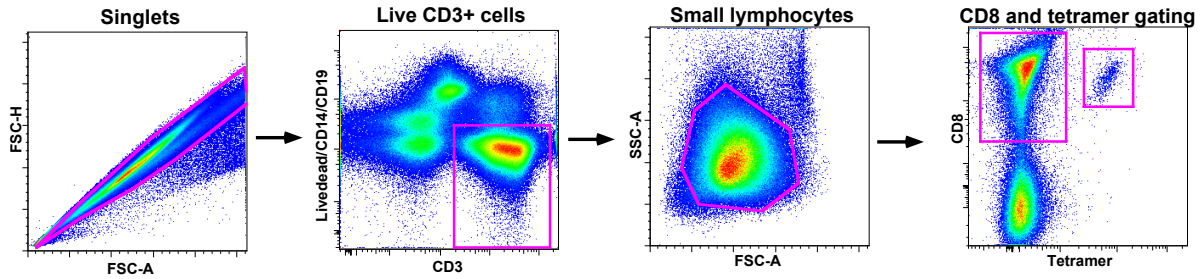


Figure 2.3a. Example of flow cytometric gating strategy used to identify CD8+ and tetramer-positive cell populations.

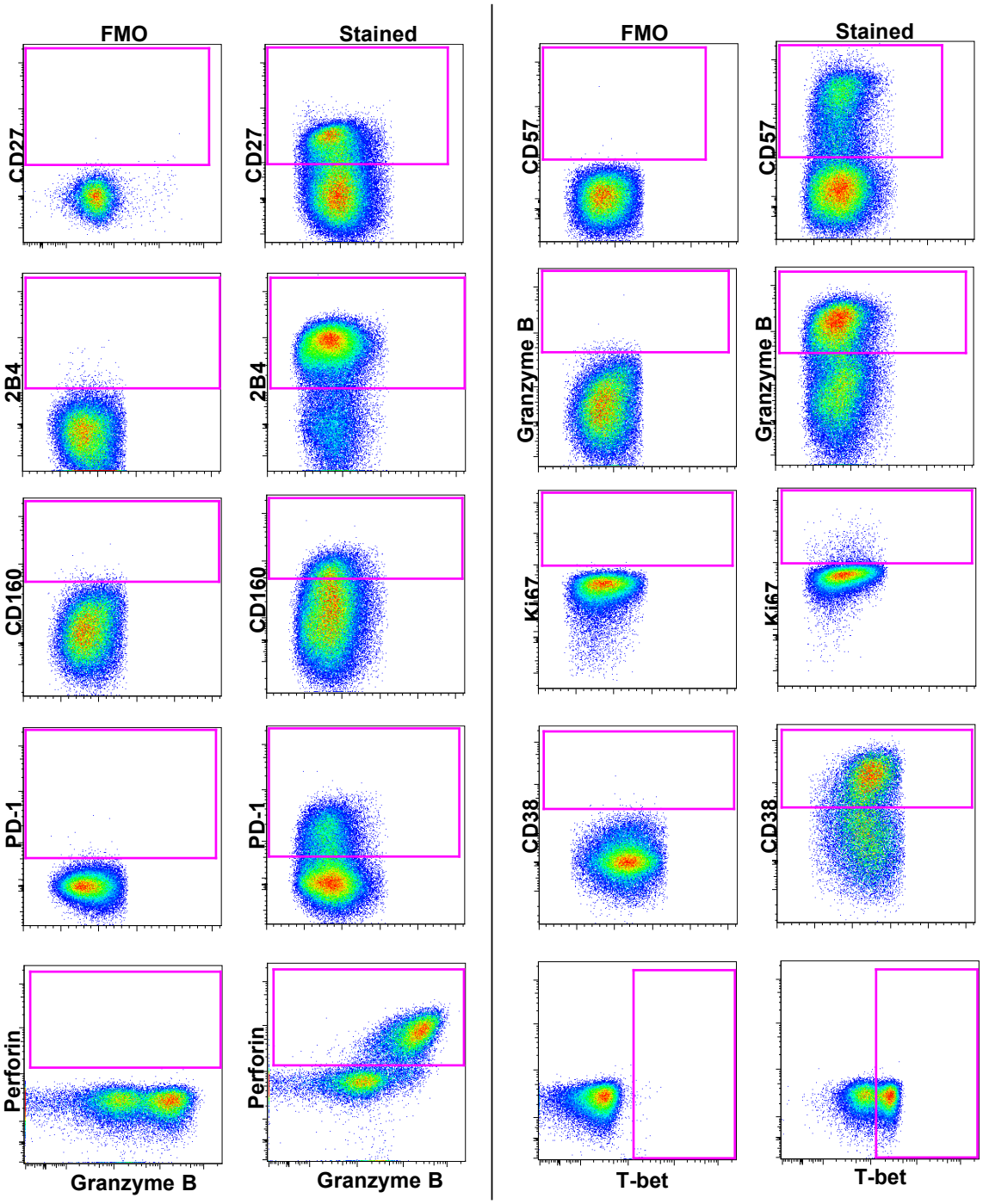


Figure 2.3b. Examples of flow cytometric gating used during tetramer staining experiments, using fluorescence-minus-one (FMO) controls. Shown are gates on CD8+ T-cell populations, which were subsequently applied to tetramer-positive populations within the same subjects.

2.23 Statistical analyses

All statistical analyses and graphical presentations were performed using Microsoft Excel for Macintosh 2011 and Graphpad Prism version 5.0 for Macintosh [202]. Comparisons between variables were performed using the Mann-Whitney U test (two-tailed), Kruskal-Wallis test with Dunn's post-test for multiple comparisons or the Wilcoxon matched pair test (for comparisons between total CD8+ and virus-specific CD8+ T-cells made within individuals). Discrete data were compared by the chi-square test or Fisher's exact test. Correlations between non-normally distributed data were made using the Spearman's rank correlation coefficient.

CHAPTER 3 – THE MOLECULAR EPIDEMIOLOGY OF HIV-1 IN THE GAMBIA AND CAIÓ, GUINEA-BISSAU: DIFFERENCES IN SUBTYPE DISTRIBUTION, INCLUDING THE PRESENCE OF CRF49_cpx, A NOVEL CIRCULATING RECOMBINANT FORM IN THE GAMBIA.

3.1 INTRODUCTION

The high heterogeneity of HIV-1 group M is due to a number of biological characteristics, including rapid viral turnover rates and an error-prone reverse transcriptase lacking the ability to proof-read [23-25]. Consequently, the HIV-1 global epidemic has seen diversification of group M into nine subtypes (A – D, F – H, J and K) and (prior to the current study) 48 CRFs [27]; the latter being strains composed of genome segments from at least two different subtypes, along with isolation from three or more epidemiologically unlinked individuals. Inter-subtype differences at the amino acid level vary between 10% in *pol* and 25 – 30% in *env* and distinct geographical patterns in HIV-1 subtype distribution can be observed [28, 203]. Whereas some epidemics, such as that in southern Africa, are almost exclusively due to a single subtype (subtype C), HIV-1 infections in central African countries tend to be far more diverse [204]. In fact, most subtypes and CRFs have been reported in Cameroon and the Democratic Republic of the Congo (DRC) [205-207], with the greatest group M diversification seen in the Congo river basin – thought to be the site of initial cross-species transmissions from SIVcpz in chimpanzees to HIV-1 in humans [207, 208]. Although the HIV-1 diversity in west Africa is still higher than in southern or east Africa, a gradual reduction in the variety of circulating subtypes and CRFs is observed as one moves further west

from Cameroon towards Ghana, Senegal and the Gambia; with CRF02_AG said to be increasingly dominant [28].

Reliable data on the HIV epidemic in the Gambia are currently lacking. The most recent national prevalence data from sentinel surveys in 2006 suggest an overall rate of 2.8% in the Gambian general population, which may represent an upward trend [209]. The most recent published data from the MRC Laboratories GUM clinic indicate that although HIV-2 infection frequency is declining in patients attending the clinic, the HIV-1 prevalence rose from 4.2% in 1988 to 17.5% in 2003 [210]. Information on the genetic diversity of the local HIV-1 subtypes is also not abundant. Prior to the current studies, the Los Alamos HIV Database (LAHDB) [27] listed only 31 sequence entries reporting subtype information from the Gambia, while the surrounding country Senegal has 895 reports, neighbouring Mali has 392, and Guinea-Bissau has 115 (following exclusion of multiple sequences from the same subject).

The prevalence and incidence of HIV-1 in Guinea-Bissau has also risen during the last two decades. Whereas HIV-1 was rare in the rural community cohort in Caió in 1990 (prevalence 0.5%), the most recent sero-survey in 2007 revealed an HIV-1 prevalence of 3.6% [51]. A cross-sectional population survey in the capital Bissau showed an HIV-1 prevalence of 4.6% in 2006 [52]. Despite these more comprehensive recent epidemiological data from Guinea-Bissau (when compared to the Gambia), there are very few studies to date describing the nature of circulating HIV-1 strains from this region [211, 212]. The limited data that are available suggest that the HIV-1 epidemic is dominated by subtype A and CRF02_AG.

Detailed sequence data are required to document the AIDS epidemic correctly, to trace the infection history, monitor changes in infection patterns and to maintain sensitive and accurate viral diagnostics. Exploration of local HIV-1 molecular epidemiology alongside regional and global sequence data can provide insight into routes of HIV-1 importation and transmission within a country; which in turn can inform public health preventative strategies. In addition, several recent studies suggest that differences in disease outcomes and response to antiretroviral therapy may exist between HIV-1 subtypes [213-215]. Furthermore, whether future HIV-1 vaccine strategy is based on immunogens optimised for local strains, or recently described 'global' mosaic vaccines that maximise coverage across HIV-1 strains worldwide [216, 217], ongoing documentation of HIV-1 sequence diversity is crucial. The current study represents an attempt to improve the local HIV-1 sequence database.

3.2 METHODS

Laboratory methods used are detailed in Chapter 2, sections 2.2 – 2.4 and 2.7 – 2.8.

Subject selection

The Gambia

HIV-1 subtyping studies were conducted in two phases. First, an exploratory study was undertaken using full-length HIV-1 *env* gp160 for subtyping, and archived plasma samples from patients attending the MRC Gambia GUM clinic. The envelope region was initially selected for amplification due to the high diversity in this gene, which allows the best discrimination of HIV-1 subtype based on a single genetic region. A group of 53 individuals were randomly selected from all those enrolled in

the cohort with a CD4 count of ≥ 28% at diagnosis (these criteria were applied in order to use the amplified products for a concurrent study). Of these, sequences from 33 individuals were successfully obtained. HIV-1 gp160 sequences from five more subjects who were recently diagnosed with HIV-1 and started ART were also obtained.

A more extensive study was then conducted in the same cohort, using *pol* sequences. The polymerase gene was chosen to maximise the ability to identify the novel CRF documented in the first phase, which displayed a distinct mosaic pattern in this region (see below). Amplification of HIV-1 *pol* from 225 randomly selected (HIV-1 or HIV-1/HIV-2 dual infected) patients with archived plasma samples was attempted. Next, PCR was attempted on the most recently diagnosed 125 HIV-1 or HIV-1/HIV-2 dually infected attendees at the MRC Gambia GUM clinic, in order to explore any recent changes in subtype distribution (as the cohort spanned approximately two decades, 1991 – 2010). Of these, sequences from a total of 294 subjects were available for subtype analysis after PCR and sequencing failures were excluded. All plasma samples used were those collected at patient enrollment into the cohort. All patients were thought to be ART naïve at this point in time.

Caió community cohort, Guinea-Bissau

HIV-1 *pol* amplification was attempted on all available subject plasma samples collected from HIV-1 or HIV-1/HIV-2 dually infected subjects during studies in 1996, 1997, 2003, 2006/7 and 2010 (see Chapter 2, section 2.1). Where multiple time points were available, the earliest sample was included in the study. A total of 157 subjects were included and sequences were obtained from 144.

HIV-1 subtyping and phylogenetic analyses on gp160 env isolates from the Gambia

HIV-1 subtype was assigned to each completed *env* sequence in the following manner. Envelope DNA sequences from each subject, along with the HIV-1 subtype reference set (2005) obtained from the LAHDB [27], additional CRF02_AG sequences DJ263 (Djibouti), MP1211 (Senegal), MP1213 (Senegal) (accession numbers AB485634, AJ251056 and AJ251057 respectively) and additional subtype A3 *env* sequences from Senegal (DD1579, DDJ360, DDJ362 and DDJ364 ; accession numbers AY521629, AY521630, AY521632 and AY521633 respectively) [218, 219] were aligned using CLUSTALW2 [220, 221]. All alignments were inspected and edited manually using Se-Al (Sequence Alignment editor, [222]), and ambiguous regions with multiple insertions and deletions were removed. Phylogenetic trees were constructed with the program PAUP* version 4.0b10 [223] using a maximum likelihood (ML) approach [224]. The trees were reconstructed under the General Time Reversible (GTR) model of nucleotide substitution [225], with proportion of invariable sites and substitution rate heterogeneity (GTR+I+G). The statistical robustness of the ML topologies was assessed by bootstrapping with 1000 replicates using the neighbour-joining method.

Phylogenies of env, p24 and near full-length sequences from CRF49_cpx isolates

Env fragments from six individuals designated as subtype J-like using the phylogenetic analyses described above were further aligned with all available subtype J *env* sequence of approximately 1200bp or above in length in the LAHDB: SE92809 (AF082394), SE9173 (AF082395), MBTB4 (AJ401046), KTB147 (AJ401041), MBS41 (AJ4010145), VLGCJ1 (AY669766), VLGCJ2 (AY669767), 98BW21.17 (AF192135), GM4 (U33099), GMB22 (AJ276694) and GMB24

(AJ276695). All sequences were trimmed to the length of the shortest sequence, thus an alignment containing 1125bp fragments were used to build a subtype J *env* phylogenetic tree using the methodology described above.

The p24 sequence from these six individuals were also aligned with HIV-1 subtype A and CRF02_AG reference isolates from the LAHDB (2005) subtype reference set [27], additional *gag* sequence from three CRF02_AG isolates SE7812 (AF107770), MP1211 (AJ251056), MP1213 (AJ251057), three A3 Senegalese isolates DDJ360 (AY521630), DDI579 (AY521629), DDJ369 (AY521631) [218, 219], additional subtype A1 isolates SE7535 (AF069671), SE8891 (AF069673), SE8131 (AF107771), SE8538 (AF069669) and the DRC isolates MBTB4 (AJ404293), KCC2 (AM000053), KTB13 (AM000054) and KTB035 (AM000055). A phylogenetic tree was reconstructed with the methodology described above.

Near full-genome sequences obtained from three of these individuals were aligned with the 2005 LAHDB subtype reference set and isolates 98BW21.17 (AF192135), DDJ360 (AY521630), DDI579 (AY521629) and DDJ369 (AY521631). Bayesian Markov chain Monte Carlo (MCMC) phylogenies were estimated under the GTR+I+G model of nucleotide substitution, using the program MRBAYES version 3.1.2 [226, 227]. The Bayesian MCMC search was set to 1,500,000 iterations with trees sampled every 100th generation. A maximum clade credibility tree (MCCT) was selected from the sampled posterior distribution with the programme TreeAnnotator version 1.5.2 [228], after discarding trees corresponding to a 10% burn-in. Convergence was assessed by calculating the effective sampling size (ESS) using Tracer v1.5 [229] and the run extended until all parameter estimates showed

ESS values of at least 200. The MCCT was edited with the programme FigTree version 1.1.2 [230].

Characterisation of subtype recombination in CRF49_cpx

Simplot and bootscan analyses of near full-genome isolates N18380_GM, N26677_GM and N28353_GM were performed using Simplot [231]. Pure subtypes A through K were included (and in a second analysis, isolate 98BW21.17 added) and the alignment was globally gap stripped. Sliding window was set to 400bp and increments set to 50bp. Bootscanning was performed using the neighbour-joining method, using the Kimura (two-parameter) distance model and 100 bootstrap replicates for each sliding window. The transition/transversion ratio was set to 2.0. For each CRF49_cpx sequence, markers were placed at breakpoints between subtypes and an alignment of each fragment used to construct phylogenetic trees using the maximum likelihood methodology (and bootstrapping with 1000 replicates using the neighbour-joining method) described above. The HIV Sequence Locator tool at the LAHDB was used to assign HXB2 numbering to each fragment and the Recombinant HIV-1 Drawing Tool (also at the LAHDB) utilised to construct a recombinant map of CRF49_cpx representing a consensus of breakpoints across the three full genomes.

Subtype designation of Gambian HIV-1 pol sequences

Due to the large dataset, subtype designation was initially attempted using automated web-based tools REGA subtyping tool (v2) [232] and the Stanford HIV Drug Resistance Database [233]. REGA initially constructs a phylogenetic tree using the Hasegawa-Kishino-Yano (HKY) distance method in PAUP* [223] and an alignment containing group M HIV-1 pure subtypes. A second tree is then

constructed following inclusion of CRFs in the alignment. The query sequence is then examined for recombination using bootscanning (sliding window 400bp, steps of 50bp). In the Stanford Database, each sequence is compared to a list of pure subtypes, along with CRF01_AE and CRF02_AG.

Sequences where one or both algorithms could not reliably assign a subtype or there was disagreement between the two algorithms were further characterised using reconstruction of phylogenetic trees using MRBAYES version 3.2 [226, 227]. An alignment was constructed with the query sequences, the LAHDB (2005) subtype reference set [27], additional subtype A3 sequences described above, as well as the newly identified CRF49_cpx sequences. Methodology used in MRBAYES was as above, using the GTR+I+G model of nucleotide substitution, but with the Bayesian MCMC search set to 50,000,000 iterations with trees sampled every 500th generation. An MCCT was selected from the sampled posterior distribution after discarding trees corresponding to a 25% burn-in. Isolates that did not reliably cluster with reference or known subtypes were analysed further using the NCBI genotyping tool [234]. This tool slides a window along the query sequence and compares each window with subtype reference sequences using the NCBI basic local alignment search tool (BLAST) [235], allowing characterisation of recombination and detection of possible unique recombinant forms (URFs).

Subtype designation of Caió HIV-1 pol sequences

An alignment was constructed with the query sequences, the LAHDB (2005) subtype reference set [27], additional subtype A3 sequences described above, as well as newly identified CRF49_cpx sequences. Bayesian MCMC phylogenies were estimated as above using MRBAYES, using the GTR+I+G mode of nucleotide

substitution, but with the Bayesian MCMC search set to 20,000,000 iterations with trees sampled every 500th generation. An MCCT was selected from the sampled posterior distribution after discarding trees corresponding to a 25% burn-in. All isolates were reliably subtyped using this phylogeny.

Phylogenetic reconstructions of CRF02_AG and subtype C pol isolates following inclusion of external sequences in the datasets

CRF02_AG: In order to explore whether the CRF02_AG epidemics in the Gambia and Caió are largely due to local, country/village-specific transmission or if frequent introductions from external sources occur, a large dataset consisting of CRF02_AG *pol* sequences from throughout Africa was collated. The LAHDB was first searched for sequences from the same region as the *pol* fragment amplified in this study, limiting the search to only entries from Africa classified as subtype CRF02_AG, with known sampling country and date. Multiple sequences from the same individuals and those without open reading frames were excluded. The alignment was trimmed to allow inclusion of the maximum number of sequences in the analysis (fragment of 719bp corresponding to HIV-1_{HXB2} position 2015 – 2733, reverse transcriptase codons 1 – 212). The final alignment contained a total of 1085 CRF02_AG sequences: 127 from the Gambia, 121 from Caió, Guinea-Bissau and 837 from other African countries including Senegal (129), Mali (204), Ghana (39), Burkina Faso (81), Cameroon (282), DRC, Togo (39), Gabon (21), Equatorial Guinea (22), Nigeria (8) and Djibouti (1). An ML tree was constructed using FastTree [236], using the GTR model of nucleotide evolution and a single rate for each site to account for varying evolutionary rates across sites (the 'CAT' approximation). FastTree is quick and more accurate than distance methods traditionally used for explorative phylogenies with large datasets [236]. The

reliability of each split in the tree is estimated using the Shimodaira-Hasegawa (SH) test on the three alternate topologies around that split, and 1000 resamples. This results in a 'local bootstrap' support value at each node [236]. Significant clusters (defined by a bootstrap value of ≥ 0.90) containing Gambia or Caió sequences were identified.

Subtype C: A similar analysis was carried out for HIV-1 subtype C isolates from the Gambia. BLAST [235] was used to identify the 10 most similar sequences to all Gambian subtype C *pol* isolates. An alignment was created using these sequences, along with *pol* fragments from all full-length subtype C isolates in the LAHDB with known country of sampling (total number of sequences n = 515). An ML tree was reconstructed using Fastree according to the parameters described above.

3.3 RESULTS

Demographics of the subjects included

The Gambia

Ten subjects were included in both the exploratory study (n = 38) and the subsequent more extensive study (n = 294). The total number, therefore, of HIV-1 infected subjects in the Gambia that were included in the final subtype assignment was 322. Median age at diagnosis was 35 (range 16 – 72), 69.8% (225/322) were female and 94.2% were HIV-1 mono-infected (remainder were HIV-1/HIV-2 dually infected). A higher percentage of women attending the GUM clinic in Gambia have been reported and may be due to changes in referral policies and sex-specific differences in health-care seeking behaviour [210]. Ethnicity data were available

on 277 individuals. Reflecting the ethnic diversity present in the Gambia, the following distribution of ethnicities were present amongst the subjects included: Fula (19.5%), Jola (20.9%), Mandinka (33.2%), Manjako (2.9%), Serehule (2.2%), Serere (4.0%), Wolof (11.2%), other (6.1%). Aside from a higher proportion of Jola and lower proportion of Serehule subjects in this cohort, this distribution was similar to the nationwide ethnic distribution reported in the Gambia 2003 national census data [237]: Fula (22%), Jola (11%), Mandinka (36%), Manjako (2%), Serehule (8%), Serere (3%), Wolof (14%), other (4%). All subjects were antiretroviral therapy naïve at time of sampling.

Caió, Guinea-Bissau

A total of 144 subjects were included in the final subtype assignment. Median age at diagnosis was 40 (range 18 – 68), 61.8% (89/144) were female and 70.2% were HIV-1 mono-infected, with the remainder being HIV-1/HIV-2 dually infected (significantly greater HIV-1/HIV-2 dually infected individuals included when compared to the MRC Gambia GUM cohort, $p < 0.0001$). All subjects were antiretroviral therapy naïve at time of sampling. The proportion of those diagnosed in surveys in 1990, 1996/7, 2003 and 2006/7 were 2.8%, 36.8%, 12.5% and 47.9% respectively.

HIV-1 subtyping using env sequences isolated from the Gambia

The subtype assignment of the initial 38 *env* sequences was obtained by aligning the sequences with LAHDB HIV-1 (2005) subtype reference sequences (which includes approximately four reference sequences from each relevant subtype), along with an additional three CRF02_AG and four A3 sequences (two from A3/CRF02_AG recombinants) as described above and constructing a maximum

likelihood tree. As none of the new Gambian *env* sequences clustered with currently known recombinant forms other than CRF02_AG, for clarity Figure 3.1 displays reference isolates from pure subtypes and CRF02_AG only. Five of the new Gambian sequences (N057856_GM, N059096_GM, N9845_GM, N75698_GM and N040736_GM) clustered with the Senegalese A3 (DDJ360, DD1579) and A3/CRF02_AG recombinant (DDJ364, DDJ362) sequences [218, 219] with a bootstrap support of 81% (see Figure 3.1 cluster denoted by ∞). Given the regional frequency of A3-like viruses, their occurrence in Gambia is not unexpected. Four isolates (N59677_GM, N058521_GM, N22314_GM and N018622_GM) clustered with reference and Gambian CRF02_AG sequences (bootstrap support 83%), although it can be difficult to distinguish subtype A (A1, A2, A3) from CRF02_AG isolates based in *env* alone as this region is largely subtype A derived in CRF02_AG [238]. An additional four isolates did not form significant clusters (N32458_GM, N47046_GM, N058628_GM and N006909_GM). From this analysis, it appears that the heterogeneity within the global CRF02_AG subgroup is equally reflected within the Gambian AG viruses. It is clear that the subtype A *env* sequences from circulating Gambian strains are distinct from both A1 and A2 reference isolates in the LAHDB, and more closely related to Senegalese A3 or CRF02_AG isolates.

In addition to the A and AG-like isolates, the novel viruses include a single subtype B (N059733_GM), three subtype C isolates (N005312_GM, N25667_GM, N025015_GM) and two subtype D isolates (N73603_GM, N001823_GM) clustering with high bootstrap values within the reference isolate clusters for these subtypes (Figure 3.1). Of special interest were six isolates (N18380_GM, N001605_GM, N24017_GM, N28353_GM, N005284_GM and N26677_GM) forming a monophyletic

cluster within the subtype J branch (bootstrap value of 100%, see Figure 3.1 and below).

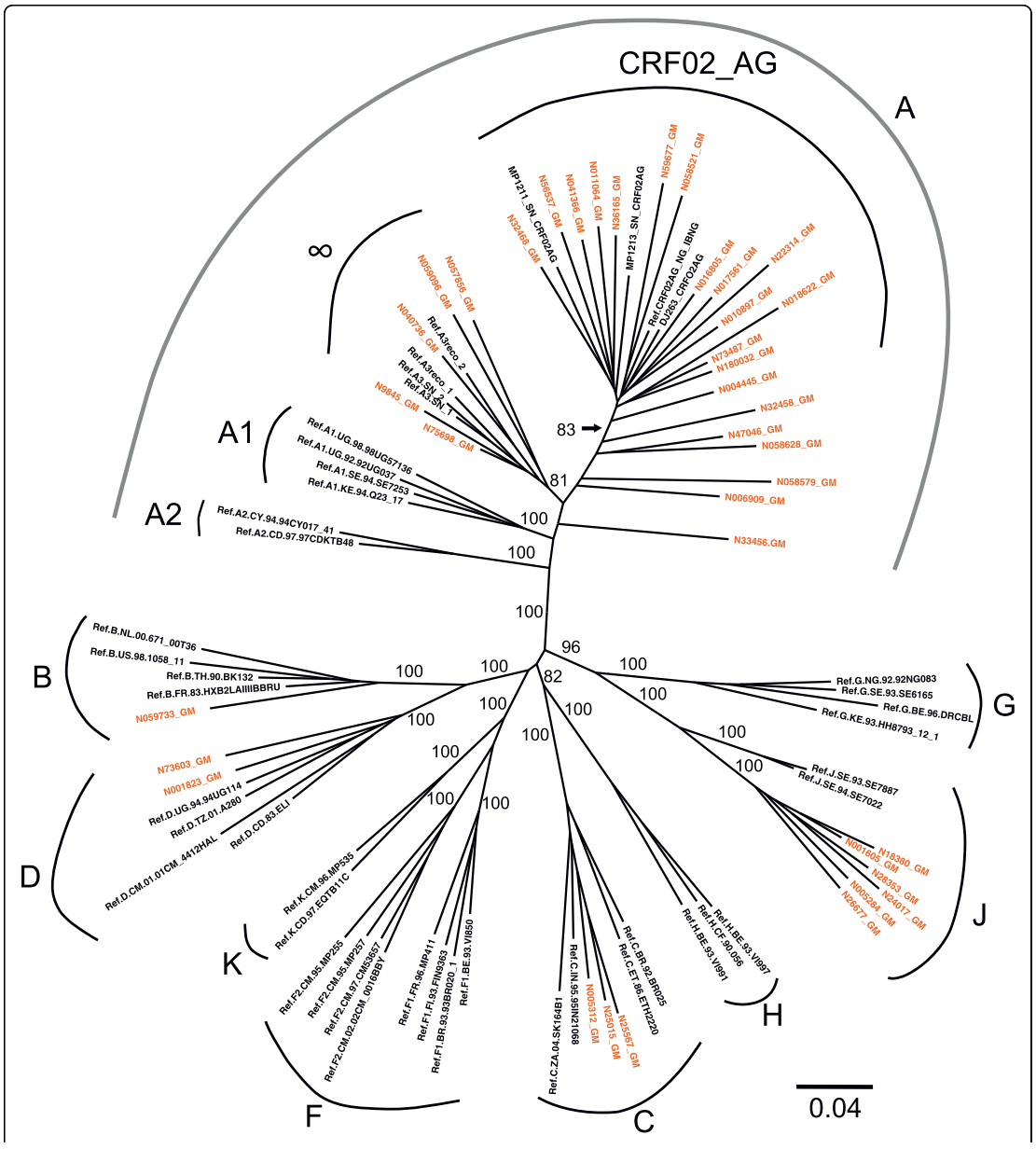


Figure 3.1 Phylogenetic classifications of 38 new Gambian HIV-1 full-length *env* sequences (highlighted in red), along with reference subtypes and additional subtype A sequences (CRF02_AG and Senegalese A3 variants). Bootstrap percentiles above 70% from 1000 replications (using the neighbour-

joining method) are shown at the corresponding branches defining major grouping of sequences. Five of the new Gambian sequences cluster with the Senegalese A3 variant sequences with a bootstrap support of 81 (∞). Branch lengths represent the number of substitutions per nucleotide sites.

Gambian isolates with subtype J-like env have subtype A gag regions

Three early Gambian HIV-1 samples, GM4 (U33099), GM5 and GM7, were reported to be distinct from the pure HIV-1 subtypes A to G known at the time [239] when the J subtype had not yet been defined. GM4 is described in the LAHDB as a subtype CGJ mosaic, although phylogenetic analyses suggest that it is subtype J-like in *env* [240]. Since that time, two additional Gambian J-like *env* sequences were reported (GMB22, GMB24 [240]). GenBank was searched for sequences with genetic similarity to either the GMB22 or the N28353 sequences and additional subtype J *env* sequences were identified: VLGC-J1 (*env* from a virus identified in Germany), VLGC-J2 (of unknown origin) [241], the 98BW21.17 isolate from Botswana [242] and the MBTB4, KTB147 and MBS41 isolates from DRC [207]. A phylogenetic tree was constructed as described above with these isolates, along with the six subtype J-like *env* samples from the current study (Figure 3.2). All nine subtype J-like *env* sequences from the Gambia form a monophyletic cluster (with a bootstrap support of 92%) and are distinct from the DRC isolates (Figure 3.2).

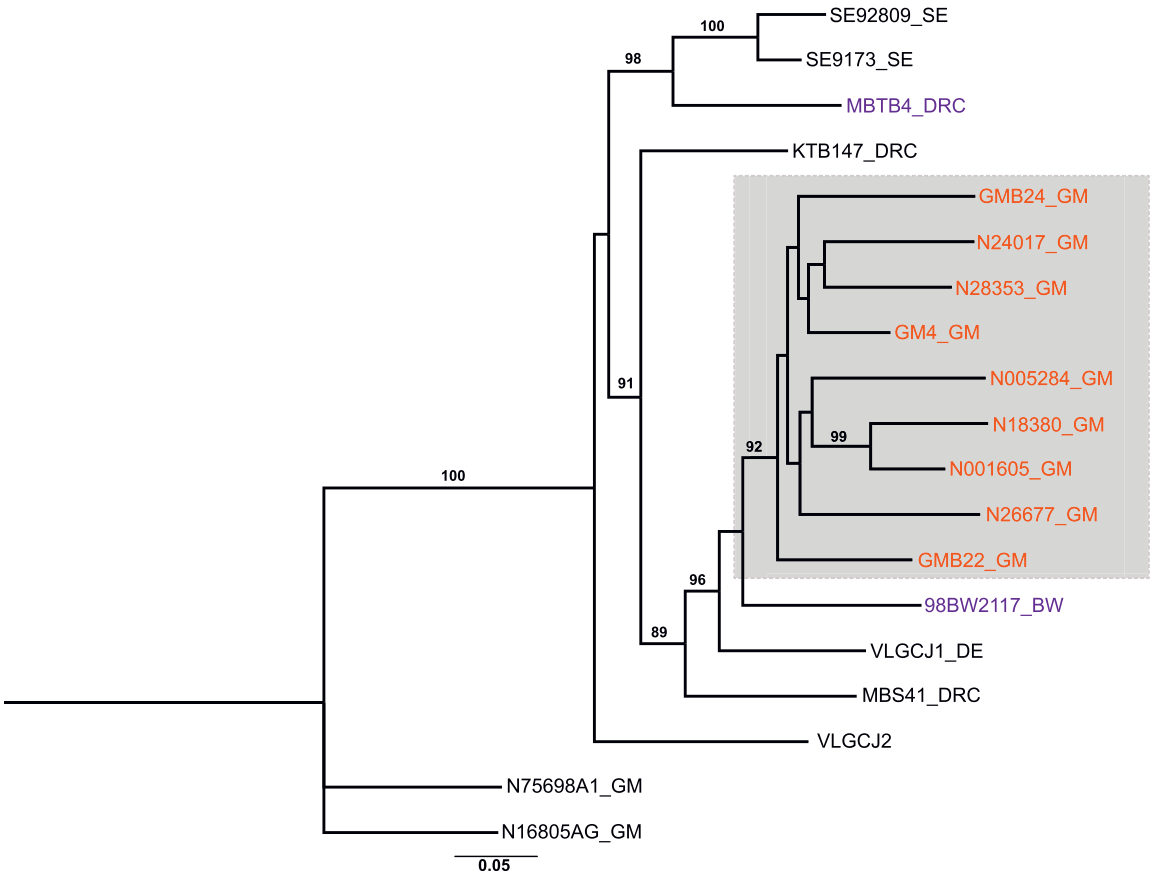


Figure 3.2 Phylogenetic tree with all available subtype J-like *env* Gambian isolates (red), including the three older isolates GM4, GM22 and GM24, and other subtype J *env* sequences from the Los Alamos HIV Database. MBTB4 and 98BW21.17 (in purple) are subtype A *gag* /J *env* recombinants described from outside the Gambia (DRC and Botswana respectively). The Gambian subtype J-like *env* monophyletic cluster is boxed. SE92809 and SE9173 are the two subtype J reference strains (From DRC, isolated in Sweden). Bootstrap percentiles above 70% from 1000 replications (using the neighbour-joining method) are shown. The tree is rooted by outgroups formed by subtype A1 and CRF02_AG *env* fragments from the Gambia (N75698A1_GM and N16805_GM). Branch lengths are expressed as the number of substitutions per nucleotide sites.

The Botswana isolate, 98BW21.17, was reported as a novel subtype A/J recombinant [242], although it has since been reclassified by the LAHDB as an AGJ recombinant, as parts of the genome are said to be more closely related to CRF06_AJGK than to any one isolate of subtype A or J [27]. The GMB22 and GMB24 isolates are also reported as having subtype A *gag* regions, although only *gag* sequence from GMB22 is available [240]. To test the idea that a novel recombinant is circulating in the Gambia, the *gag* p24 regions from the six novel J-like *env* isolates were sequenced and all were found to be subtype A. Furthermore the *gag* regions from the Botswana isolate 98BW21.17, GMB22 and five of the new A/J isolates form a monophyletic cluster with a bootstrap support of 94% (Figure 3.3). These *gag* isolates are distinct from sub-subtype A1, A2, A3 sequences, as well as those derived from CRF02_AG isolates. One new recombinant isolate (N5284_GM) *gag* region clustered with A3 [218, 219] isolates reported in surrounding Senegal, which may indicate further recombination between the novel recombinant with circulating local A3 strains. One additional isolate described in the literature, MBTB4 from DRC, is reported to have a subtype A *gag* and subtype J *env* region [207]. The subtype A *gag* phylogenetic tree was re-built including this isolate, along with three further DRC subtype A sequences (KCC2, KTB13 and KTB035), which required use of a shorter fragment length as described above. The MTBT4 isolate *gag* appears to be more closely related to subtype A *gag* regions from *gag* A/*env* J-like recombinants than other subtype A sequences (with a bootstrap support of 76%), including those from DRC (Figure 3.3). Of note, the *env* region from MTBT4 clusters with the two reference J *env*s SE9173 (from an individual known to be infected in DRC) and SE92809 (bootstrap support of 98%), rather than the other *env* J isolates with subtype A *gag* regions (Figure 3.2).

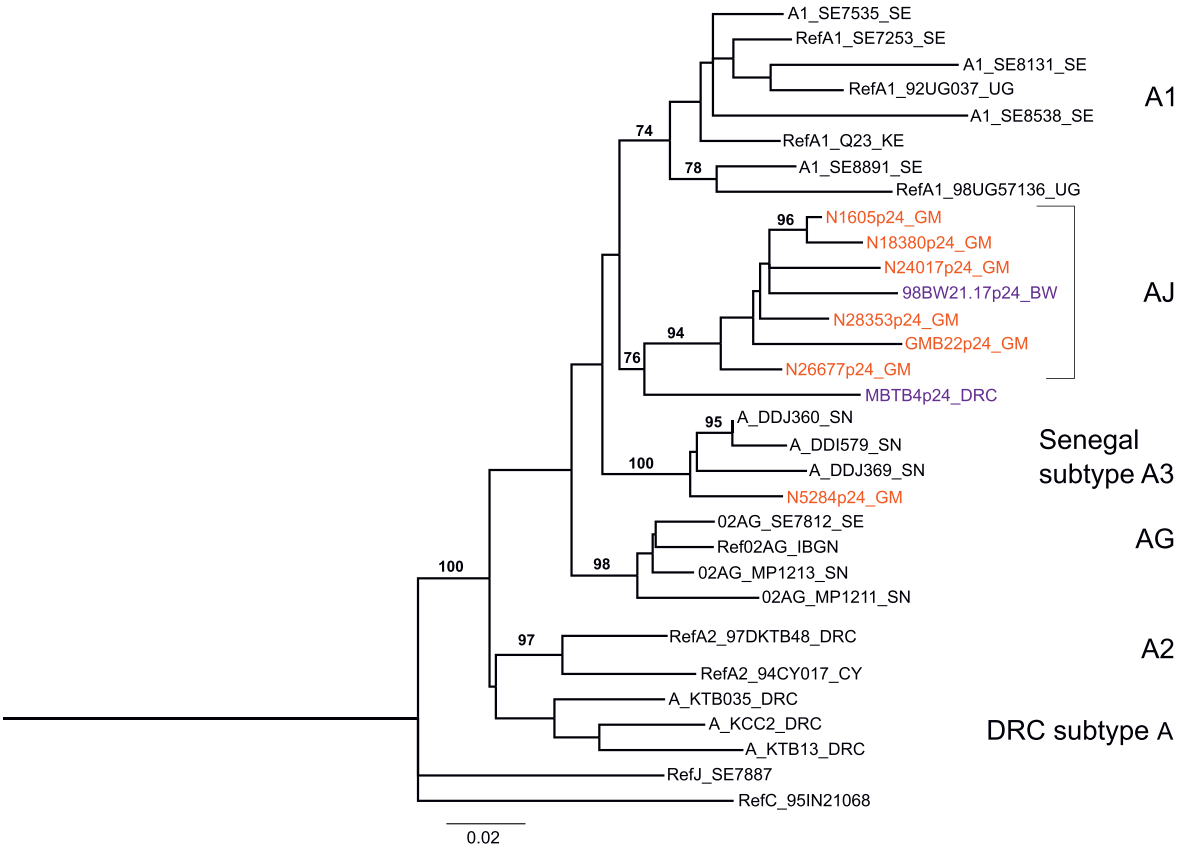


Figure 3.3 Phylogenetic tree constructed using alignments of *gag* sequence from subtype A reference strains (denoted by prefix 'Ref'), additional subtype A1 isolates, A3 isolates from Senegal, CRF02_AG isolates and subtype A *gag* sequence from isolates with subtype J-like *env* regions. Gambian isolates are in red, which includes an older isolate GM22. Sequence from the non-Gambian *gag*A/envJ recombinants 98BW21.17 and MTBT4 are highlighted in purple. The cluster formed by *gag* A sequence from isolates with J-like *env* regions is boxed. Bootstrap percentiles above 70% from 1000 replications (using the neighbour-joining method) are shown. The trees are rooted by outgroups formed by subtype J and C reference isolates from the Los Alamos HIV Database (2005) subtype reference set (SE7887 and 95IN21068). Branch lengths represent the number of substitutions per nucleotide sites.

CRF49_cpx, a novel circulating recombinant form in the Gambia

Near full-genome sequences from three of the *gag* A/*env* J-like isolates (N18380_GM, N28353_GM and N26677_GM) were generated and a phylogenetic tree constructed (Figure 3.4), which provided confirmation that these viruses represent a novel CRF, now named CRF49_cpx in the LAHDB. The three isolates clearly form a new cluster, separate form any currently known pure subtypes or recombinants (with a posterior probability of 1) and appear to be closely related to the Botswana isolate 98BW21.17.

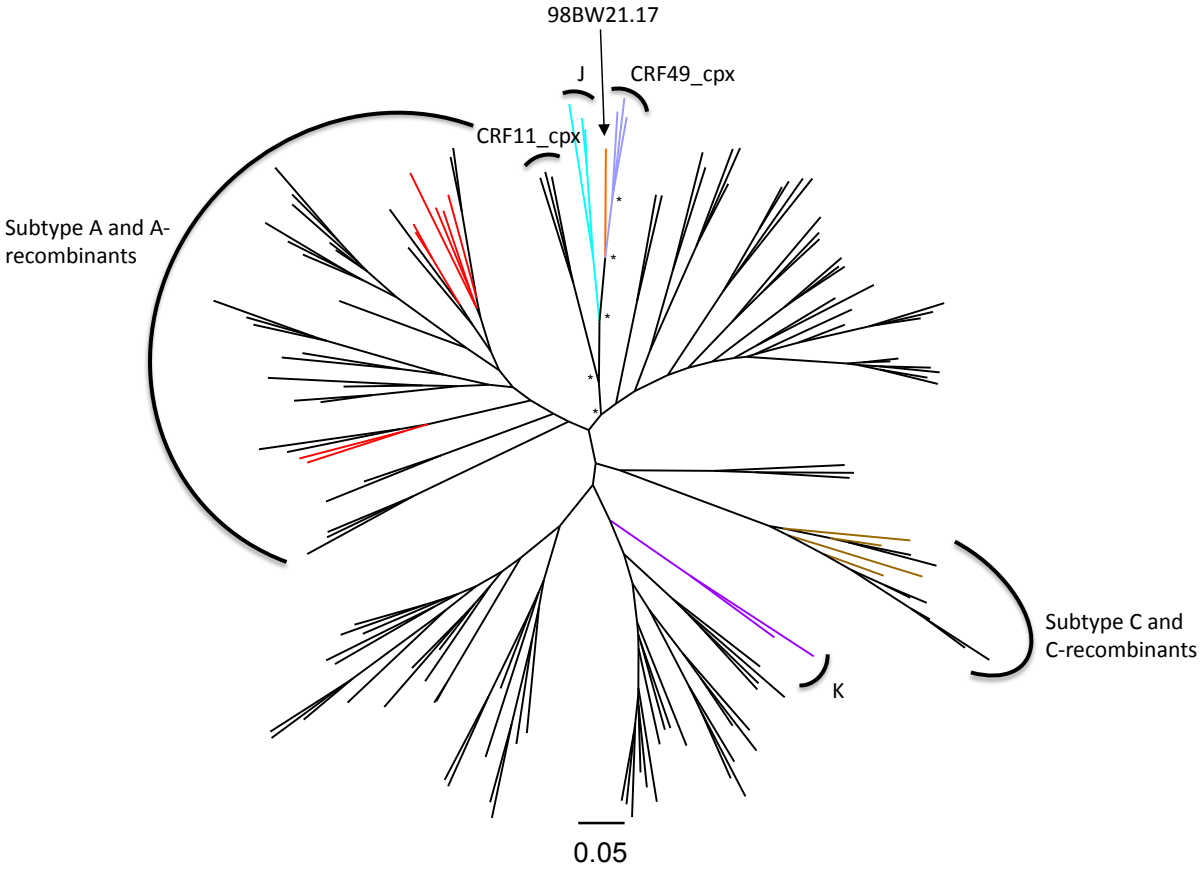
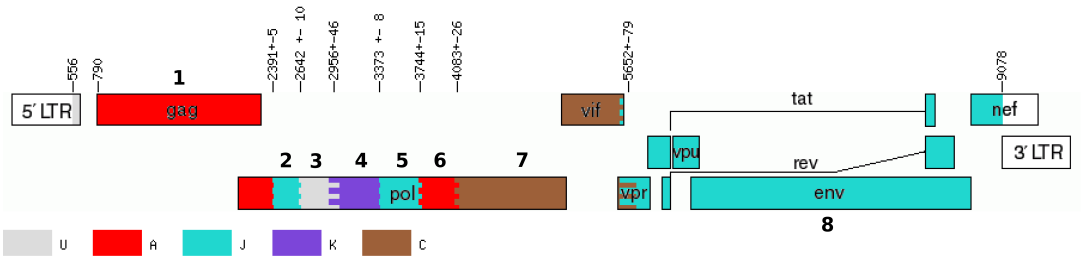


Figure 3.4 Midpoint rooted Bayesian tree using Los Alamos 2005 subtype reference set HIV-1 full genomes, additional A3 sequences, 98BW21.17 and 3

new Gambian CRF49_cpx isolates. Pure subtype sequences represented in the new Gambian complex recombinant are shown in colour (A (red), J (turquoise), C (brown), K (purple)). Relevant nodes to the new complex recombinant, with a posterior probability of 1, are marked with *.

Analyses of subtype recombination revealed a complex, but consistent pattern across the three isolates (see Figure 3.5). In addition to the largely subtype A *gag* region and J-like *env*, a significant subtype C fragment is present in a portion of *pol*, extending through *vif* to *vpr* (which is absent in 98BW21.17), where a breakpoint with the subtype J-like fragment is found. The *pol* gene is mosaic and contains regions with similarity to subtypes A, J, K and C, as well a fragment which is not clearly defined by currently known pure subtype sequences. A phylogenetic tree constructed with this fragment (not resolved through Simplot bootscanning analysis), suggested that this region was subtype F-like (Figure 3.5).

(a)



(b)

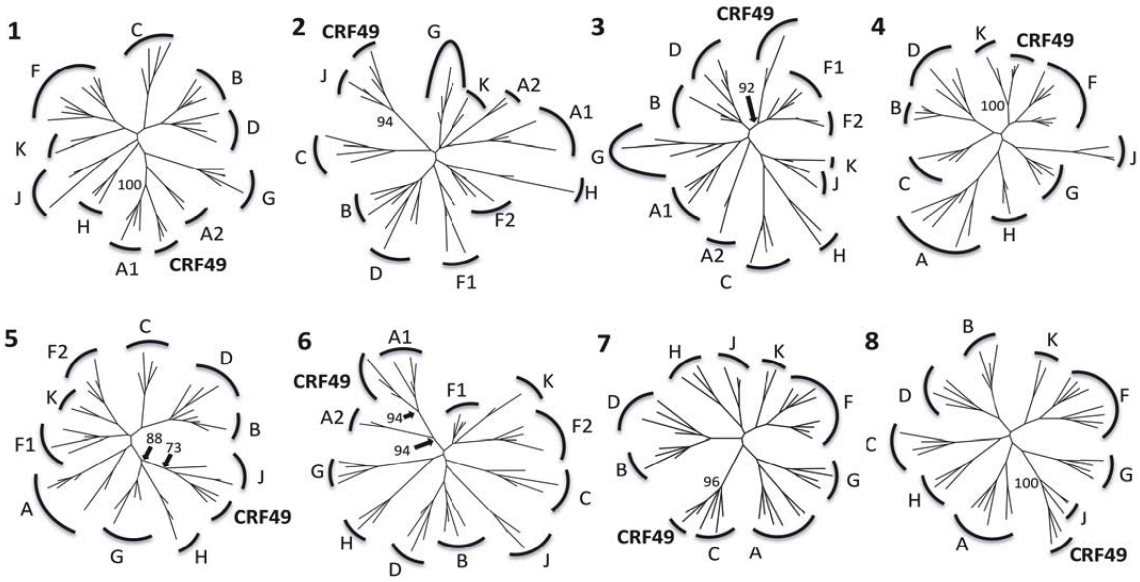


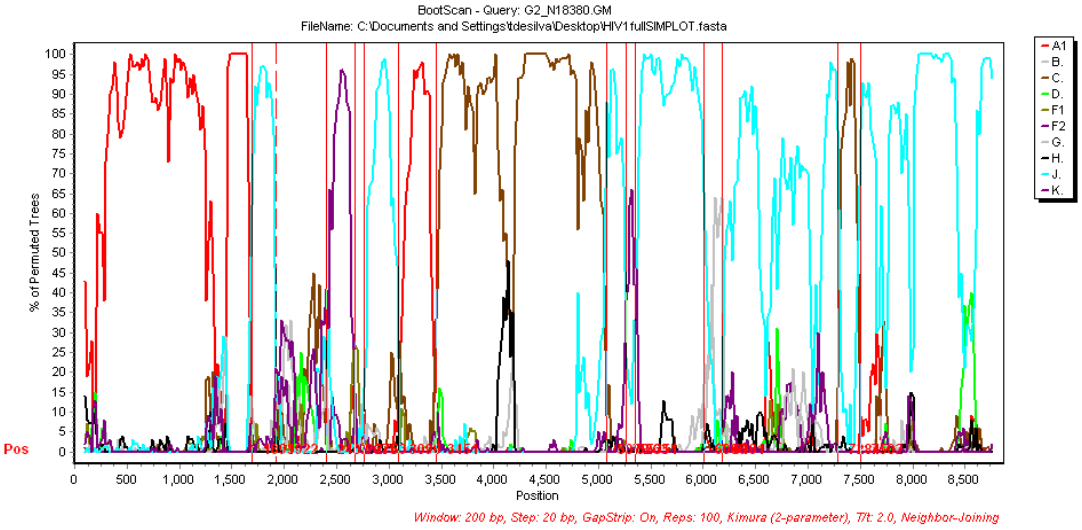
Figure 3.5 Recombinant map of CRF49_cpx and phylogenetic trees constructed from each non-recombinant fragment. (a) Recombinant map of CRF49_cpx drawn using the Recombinant HIV-1 Drawing Tool at LAHDB and depicting HXB2 numbering at breakpoints estimated via bootscan analysis in Simplot [231]. (b) Maximum likelihood trees constructed with each non-recombinant fragment (1 to 8) showing relationship to pure subtypes. Only bootstrap values (from 1000 replications using the neighbour-joining method) at relevant nodes to the CRF49_cpx isolates and closest pure subtype are shown for ease of presentation.

Simplot and bootscan analysis [231] clearly showed a similar pattern of subtype recombination across the three isolates, although there was variation in where the exact breakpoints are (Figures 3.6a - c), especially in the highly mosaic *pol* gene. The Simplot and bootscan analysis [231] was repeated for each sequence, with inclusion of the Botswana isolate 98BW21.17 in the reference set. This suggested

that apart from the subtype C-like fragment, the CRF49_cpx sequences are more similar to 98BW21.17 than to most pure reference subtypes representing each recombinant fragment (Figure 3.7). It is possible, therefore, that CRF49_cpx originated via further recombination between a 98BW21.17-like strain and a subtype C isolate.

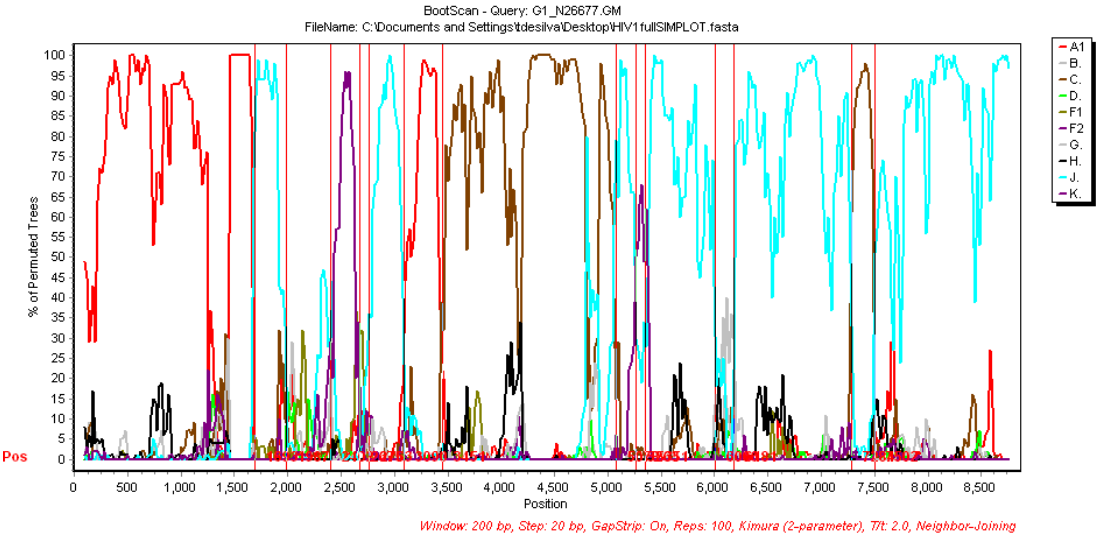
(a)

N18380.GM



(b)

N26677.GM



(c)

N28353.GM

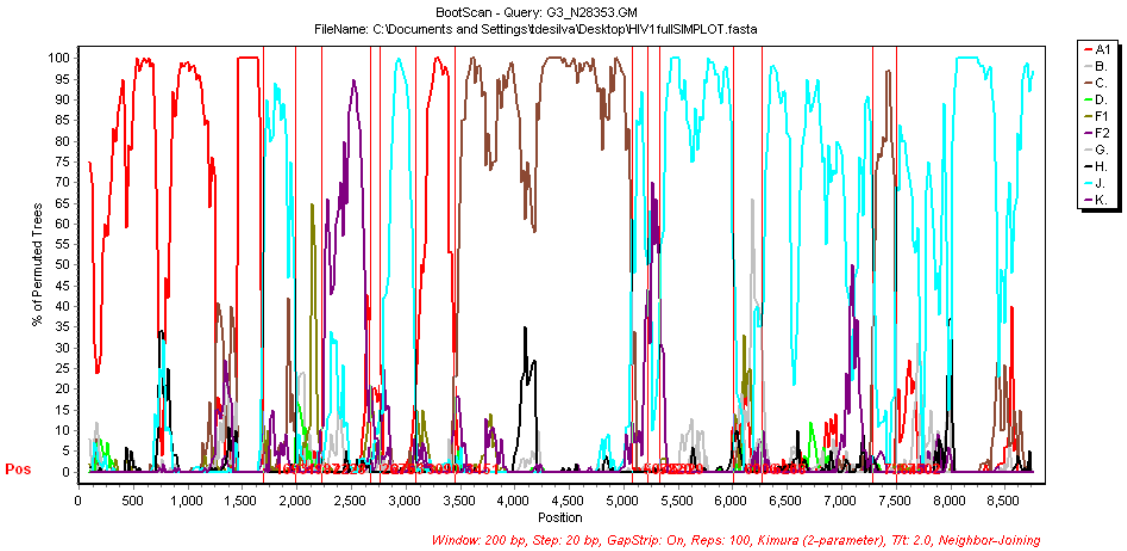


Figure 3.6 Bootscan analyses of CRF49_cpx isolates N18380_GM (a), N26677_GM (b) and N28353_GM (c) performed with Simplot [231] and including HIV-1 pure subtypes A through K.

N26677.GM bootscan with BW21.17 included

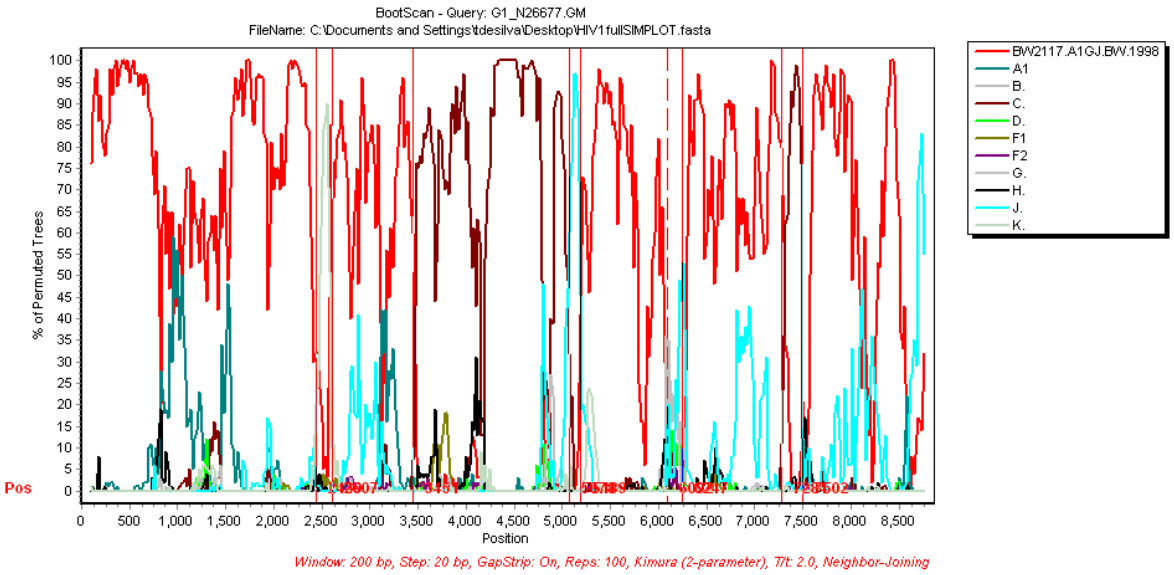


Figure 3.7 Bootscan analyses of CRF49_cpx isolate 26677_GM. Performed with Simplot [231] and including HIV-1 pure subtypes A through K and Botswana isolate 98BW21.17.

Further characterisation of HIV-1 subtypes in the Gambia

Polymerase sequence was successfully obtained from a further 294 individuals from the MRC Gambia GUM cohort. Of these 192 (65.3%) had a reliable HIV-1 subtype assigned via the REGA HIV subtyping tool and the Stanford database, which was consistent across both algorithms. A phylogenetic tree was constructed with the remaining isolates, along with the LAHDB 2005 subtype reference set and additional A3 and CRF49_cpx reference sequences (Figure 3.8). A total of 69 (67.7%) isolates clustered significantly with known reference isolates, including subtype A (9 - eight of which were A3), CRF02_AG (8), CRF06_cpx (3), CRF09_cpx (2), CRF11_cpx (2) and CRF49_cpx (45). Of the remaining sequences, 18 could be considered 'A-like' and clustered within the CRF02_AG and subtype A group.

Analysis using the NCBI genotyping tool revealed that 11 of these were likely CRF02_AG/A recombinants, although the pattern of recombination varied significantly between isolates. Other isolates were subtype D-like (3), F-like (1), G-like (3). Of the remaining 8 sequences that were still undefined (Figure 3.8), 4 were shown to be CRF02_AG/CRF09_cpx recombinants (using the NCBI genotyping tool), with a pattern of recombination suggesting these may belong to a novel CRF (Figure 3.9a - b).

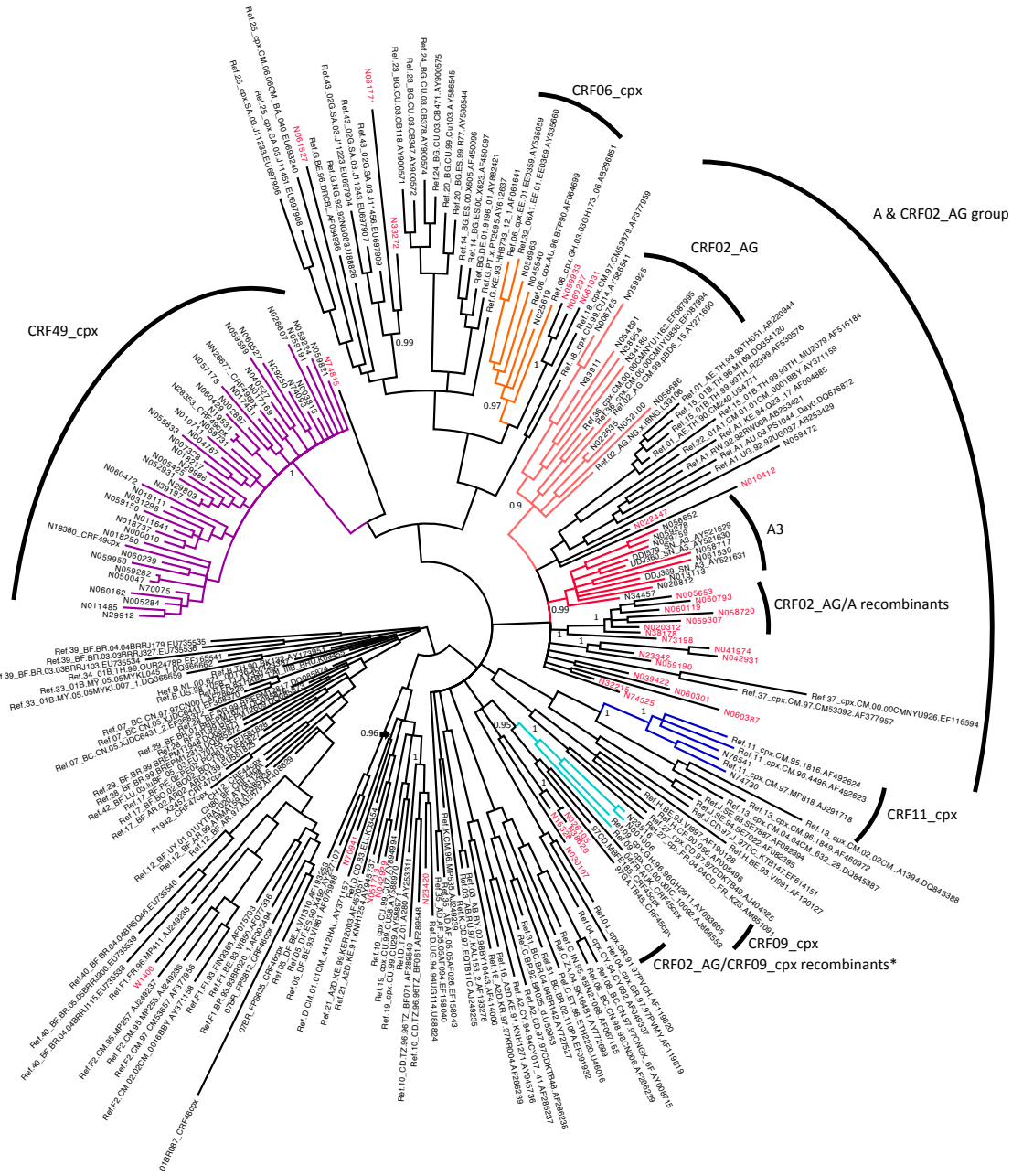
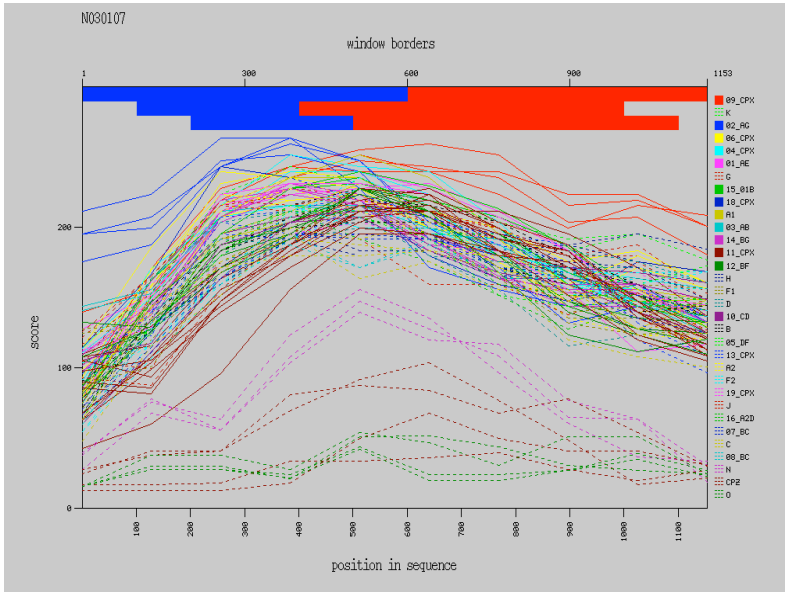


Figure 3.8 Midpoint rooted Bayesian tree using *pol* sequences not reliably subtyped with web-based algorithms REGA and the Stanford database, along with the Los Alamos 2005 subtype reference set, additional A3 sequences

and the CRF49_cpx reference isolates. Sequences that reliably clustered with known reference subtypes are shown with coloured branches for CRF49_cpx (purple), CRF06_cpx (orange), CRF02_AG (pink), subtype A3 (red), CRF11_cpx (blue) and CRF09_cpx (turquoise). Sequences that did not reliably cluster with known reference subtypes are highlighted with the sequence name in red. CRF02_AG/CRF09_cpx recombinants are denoted by *. Branch lengths and scale represent nucleotide substitutions per site. Values at relevant nodes are posterior probabilities.

(a)



(b)

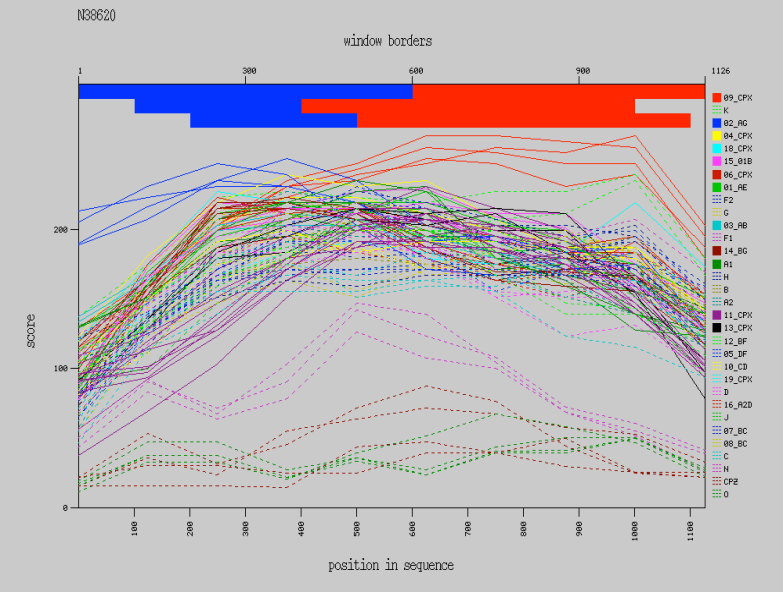


Figure 3.9 (a) and (b) Example of two CRF02_AG/CRF09_cpx recombinants. Similarity to pure subtypes and known CRFs was assessed using the NCBI genotyping tool using a sliding window approach. Both sequences show a similar pattern of recombination between CRF02_AG (blue) and CRF09_cpx (red) fragments.

Following inclusion of subjects from the initial exploratory subtyping study described above (therefore total included n = 322), the proportion of each subtype in the dataset was calculated (Figure 3. 10). Although CRF02_AG accounted for the majority of isolates (52.5%), the diversity of HIV-1 in the Gambia was found to be high, with CRF49_cpx responsible for a significant proportion of infections (14.9%).

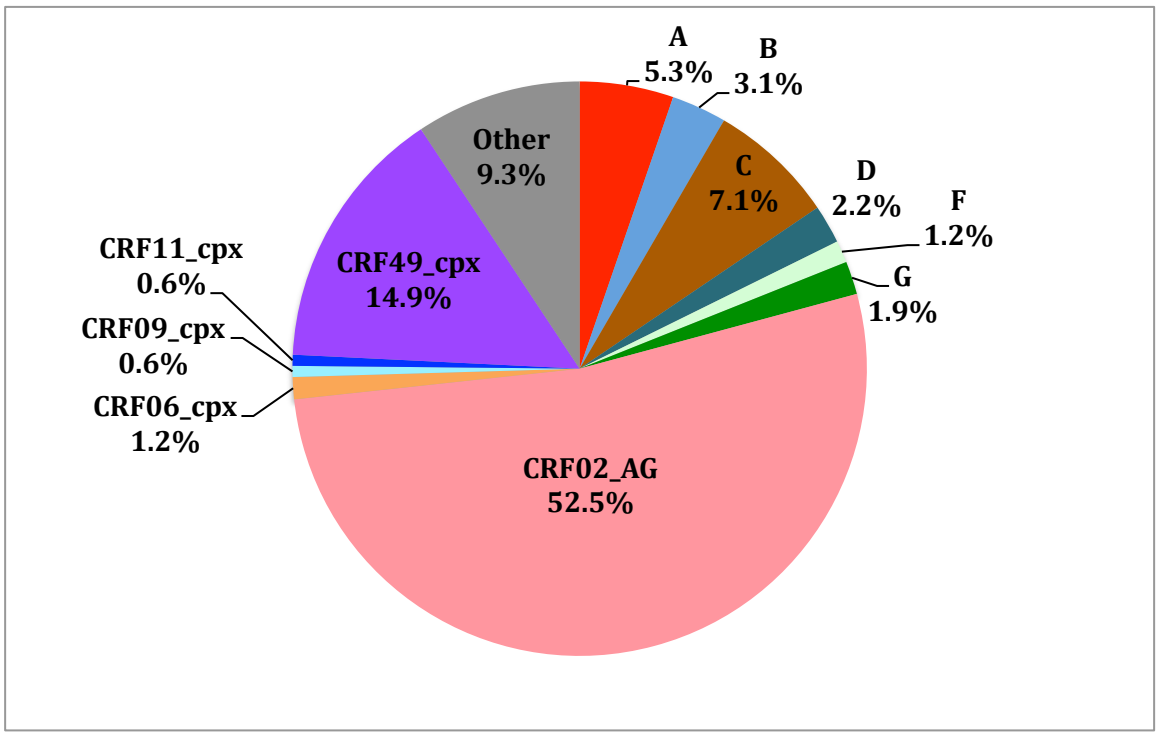
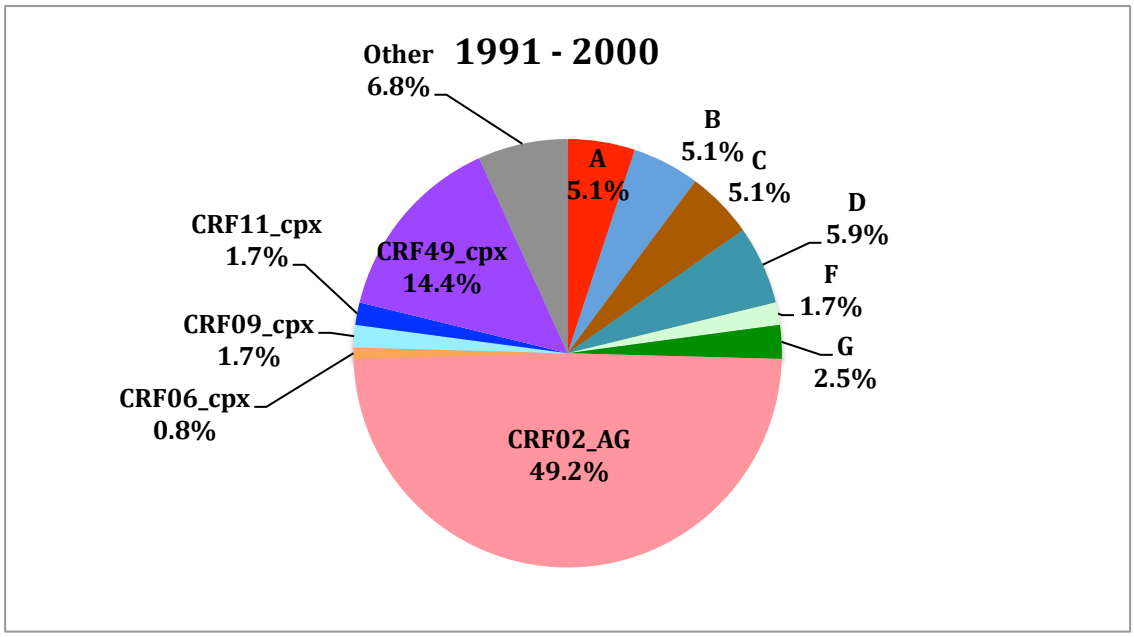


Figure 3.10 HIV-1 subtype distribution in the Gambia (n = 322). Isolates classified as A-like, D-like, F-like and G-like and included in the 'other' category. Subtype A3 isolates are included in the subtype A category.

Temporal changes in HIV-1 subtype distribution in the Gambia

Isolates included in the study were from subjects diagnosed with HIV-1 infection over approximately two decades (1991 – 2010). To see if the HIV-1 subtype distribution has altered over time, the dataset was divided into two time periods: 1991 – 2000 and 2001 – 2010, including 118 and 204 subjects in each group respectively (Figure 3.11). No significant change in either of the dominant CRFs (CFR02_AG and CRF49_cpx) was observed between the two time periods (p =0.16 and p=0.2 respectively). This suggests that CRF49_cpx continues to cause a significant proportion of HIV-1 infections in the Gambia and there are no signs that this CRF is being displaced by the dominant CRF02_AG over time. Furthermore, when considering only isolates from new HIV-1 diagnoses made over the last 2

years at the MRC Gambia GUM clinic (n = 102, March 2008 – Jan 2010), CRF49_cpx still accounted for 14.7%.



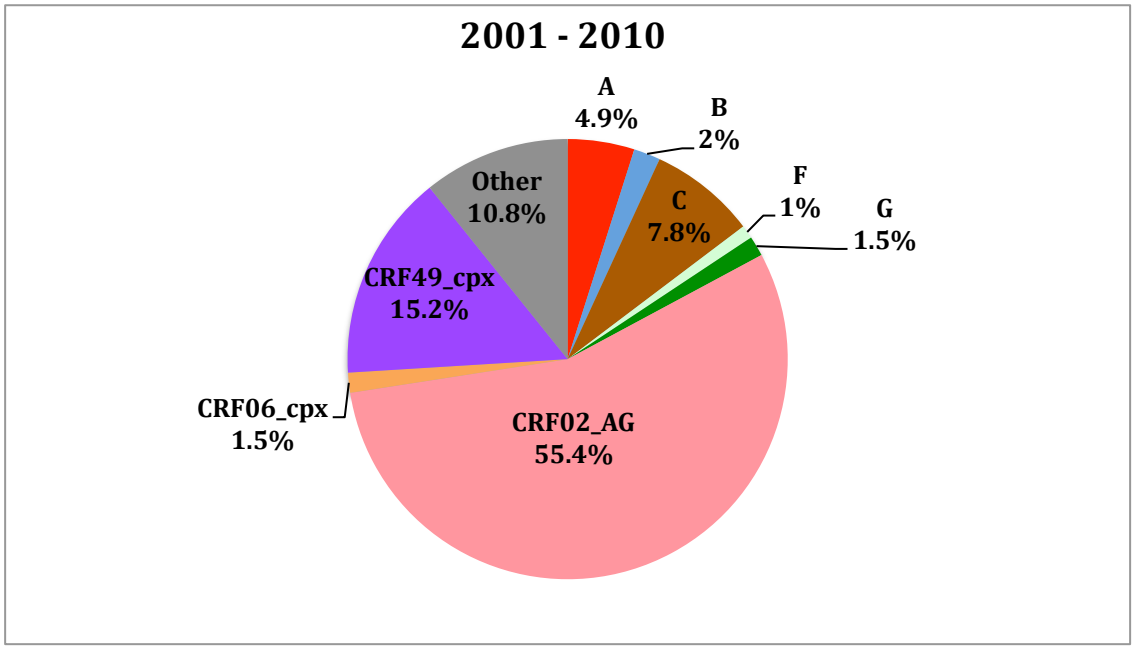


Figure 3.11 HIV-1 subtype distribution in the Gambia between 1991 – 2000 (n = 118) and 2001 – 2010 (n = 204). Isolates classified as A-like, D-like, F-like and G-like and included in the 'other' category. Subtype A3 isolates are included in the subtype A category.

HIV-1 subtype distribution in Caió, Guinea-Bissau

Polymerase sequence was successfully obtained from 144 subjects from the Caió cohort. A phylogenetic tree was constructed; along with the LAHDB 2005 subtype reference set and additional subtype A3 and CRF49_cpx reference sequences, as detailed above. CRF02_AG accounted for the majority of infections in Caió (82.9%), with subtype A (all likely A3 isolates – figure 3.12) responsible for 12.3%. Three infections (2.1%) were due to CRF06_cpx and two (1.4%) due to subtype B. The phylogenetic tree was reconstructed following exclusion of CRFs other than CRF02_AG, CRF06_cpx and CRF49_cpx for clarity (Figure 3.12).

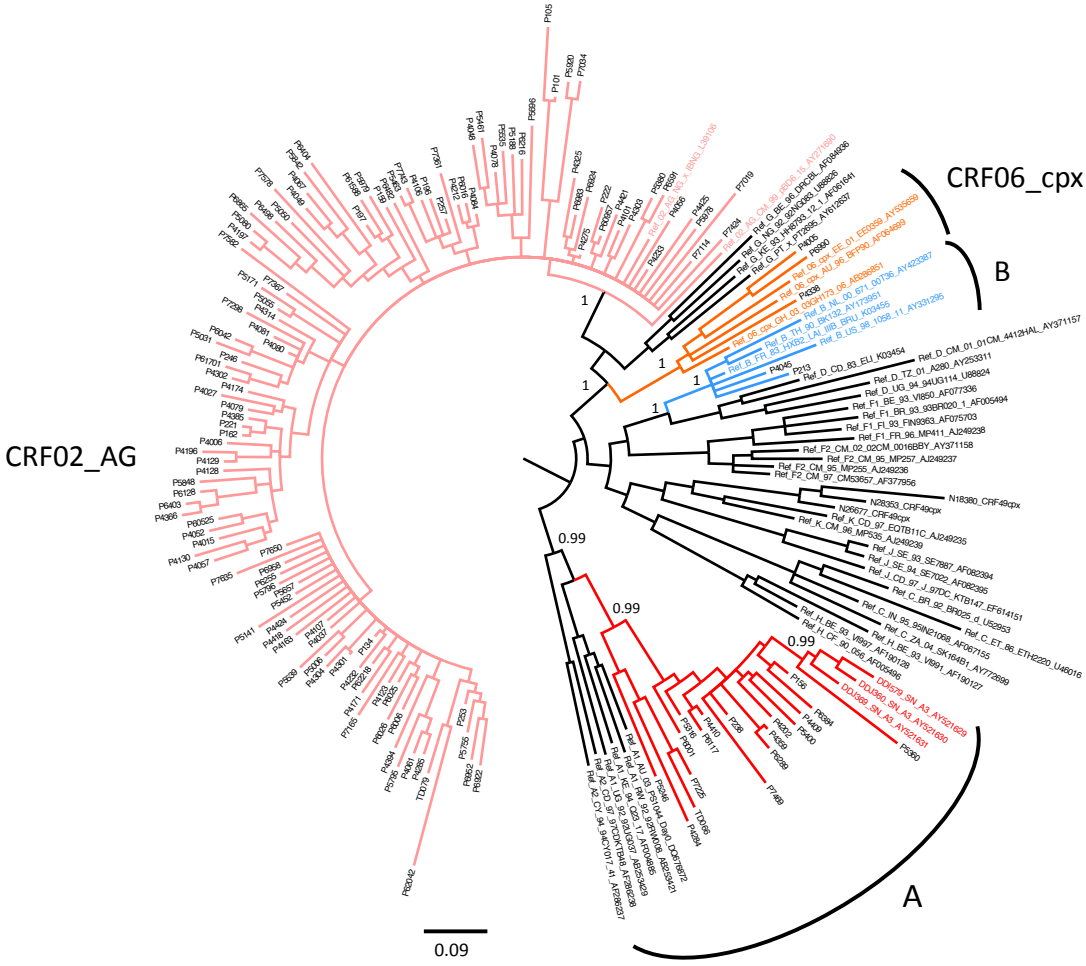


Figure 3.12 Midpoint rooted Bayesian tree using *pol* sequences from Caió, Guinea-Bissau, along with the Los Alamos 2005 subtype reference set, additional A3 sequences and the CRF49_cpx reference isolates. Sequences that reliably clustered with known reference subtypes are shown with coloured branches for subtype A (red), CRF06_cpx (orange), CRF02_AG (pink), and subtype B (blue). Reference isolate labels are also highlighted in these respective colours. Branch lengths and scale represent nucleotide substitutions per site. Values at relevant nodes are posterior probabilities.

All other available sequences (once multiple sequences from each individuals were excluded) from Guinea-Bissau available in the LAHDB were added to the dataset above, to provide an estimate of subtype distribution throughout Guinea-Bissau (Figure 3.13). It is clear that the diversity in HIV-1 sequences is greater in the Gambia than Guinea-Bissau and that CRF02_AG is responsible for a significantly higher proportion of HIV-1 infections in Guinea-Bissau ($p < 0.0001$, two-tailed Fisher's exact test).

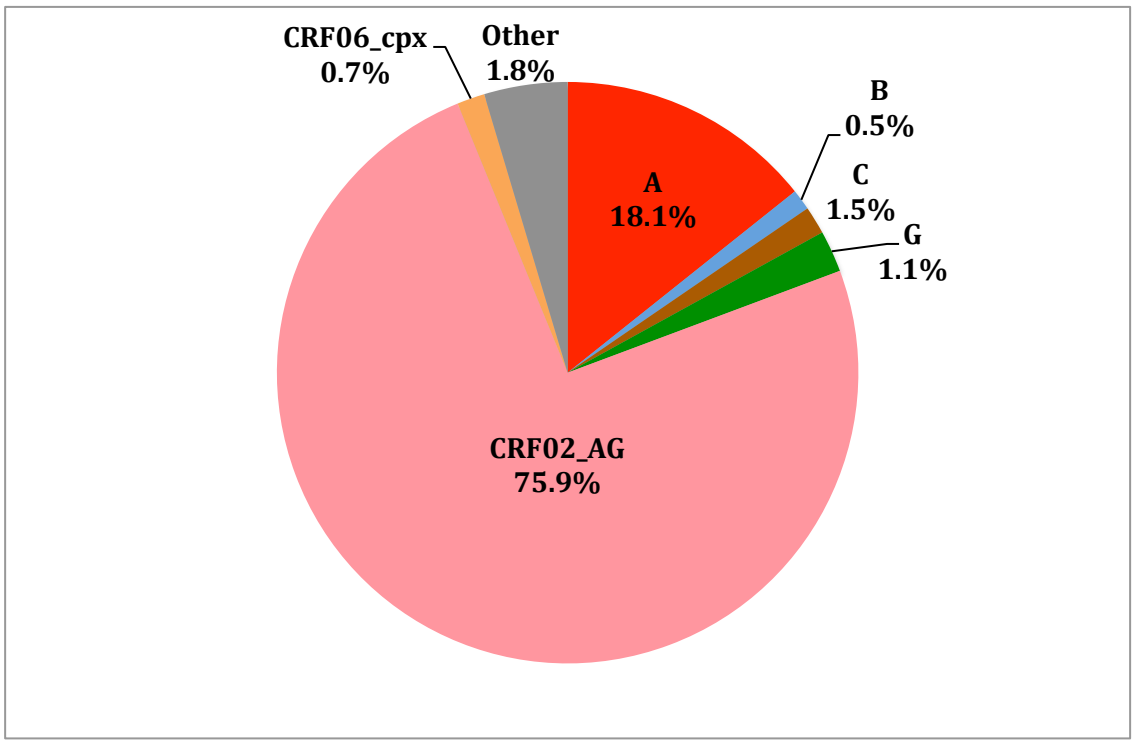


Figure 3.13 HIV-1 subtype distribution in Guinea-Bissau (n = 259, including n = 144 isolates from Caió). The 'other' category exclusively contains isolates labeled as CRF02AG/subtype A recombinants in the LAHDB.

Multiple introductions of HIV-1 CRF02_AG into the Gambia and Caió are evident

In order to explore the diversity of HIV-1 CRF02_AG isolates from the Gambia and Caió, a large dataset of 1085 African CRF02_AG samples was assembled. Isolates from the Gambia appeared in several clusters throughout the phylogeny, suggesting that multiple introductions from neighbouring countries have contributed to CRF02_AG being the dominant HIV-1 subtype in the Gambia (Figure 3.14). Gambian CRF02_AG isolates predominantly formed clusters with isolates from the neighbouring countries Senegal, Mali and Guinea-Bissau, although some were associated with isolates from more distant West African countries such as Ghana, Nigeria, Equatorial Guinea, Gabon and Burkina Faso, as well as Cameroon.

A similar pattern of multiple separate introductions was also observed in Caió, although several clusters of Caió-specific transmission were also evident (figure 3.14). Where significant clusters with external sequences were found, these were almost exclusively with Gambian sequences (with 1 Caió sequences clustering with one from Guinea Conakry).

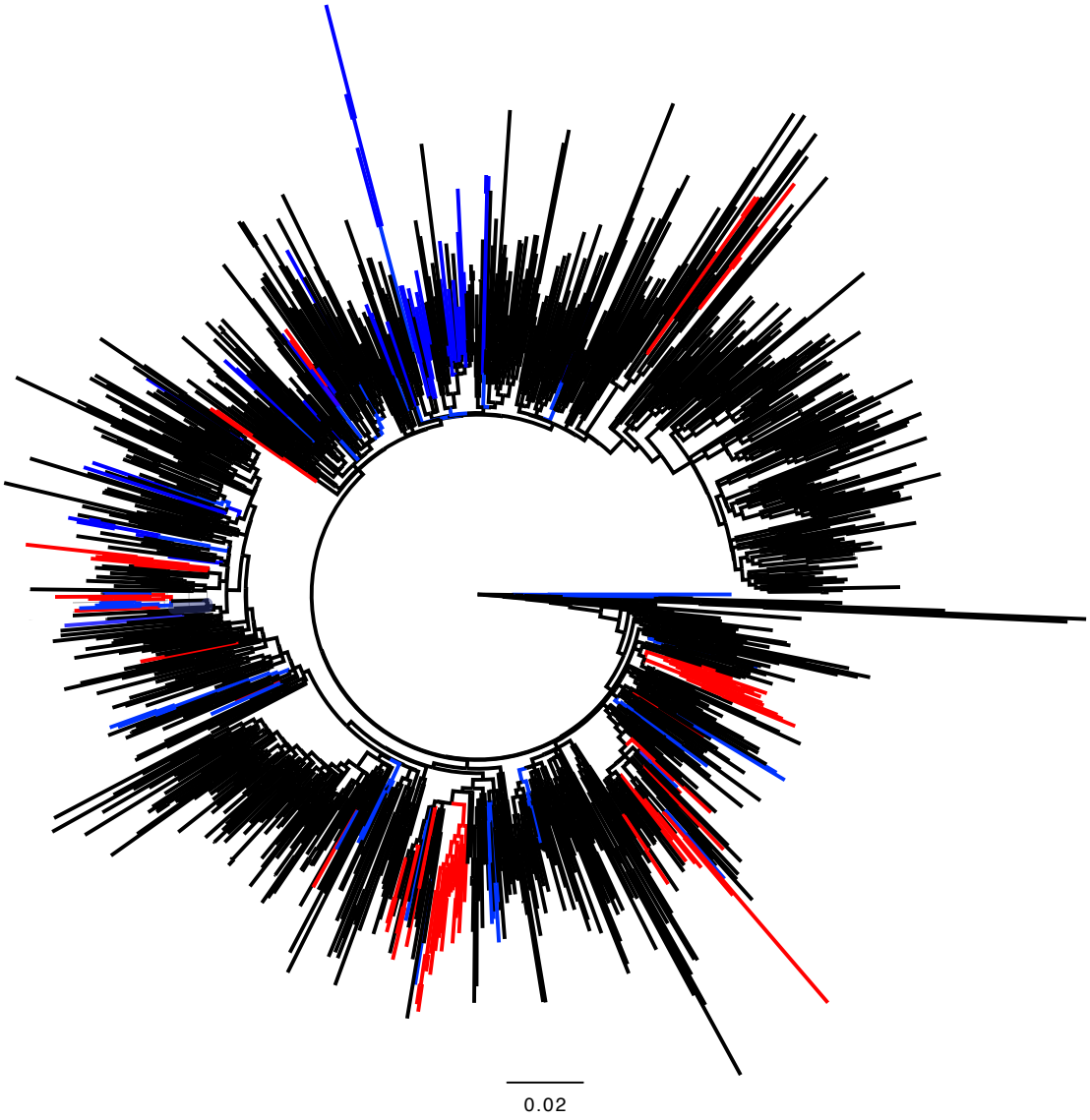


Figure 3.14 Maximum Likelihood tree using CRF02_AG *pol* sequences from West Africa (n = 1085). The phylogeny is rooted with an HIV-1 subtype B isolate. Sequences from the Gambia and Caió, Guinea-Bissau, where supported by a local bootstrap value of ≥ 0.90 at the corresponding node, are highlighted in blue and

red respectively. Branch lengths and scale represent nucleotide substitutions per site.

A single introduction may have given rise to most circulating Gambia subtype C infections

HIV-1 subtype C was the most prevalent 'pure' subtype isolated in the MRC Gambia GUM clinic cohort (7.1%). In contrast to CRF02_AG, Gambian subtype C isolates formed a monophyletic cluster in association with some isolates from Senegal (Figure 3.15). This suggests the introduction of a single lineage, followed by transmission within the sub-region, is responsible for the majority of HIV-1 subtype C infections in the Gambia.

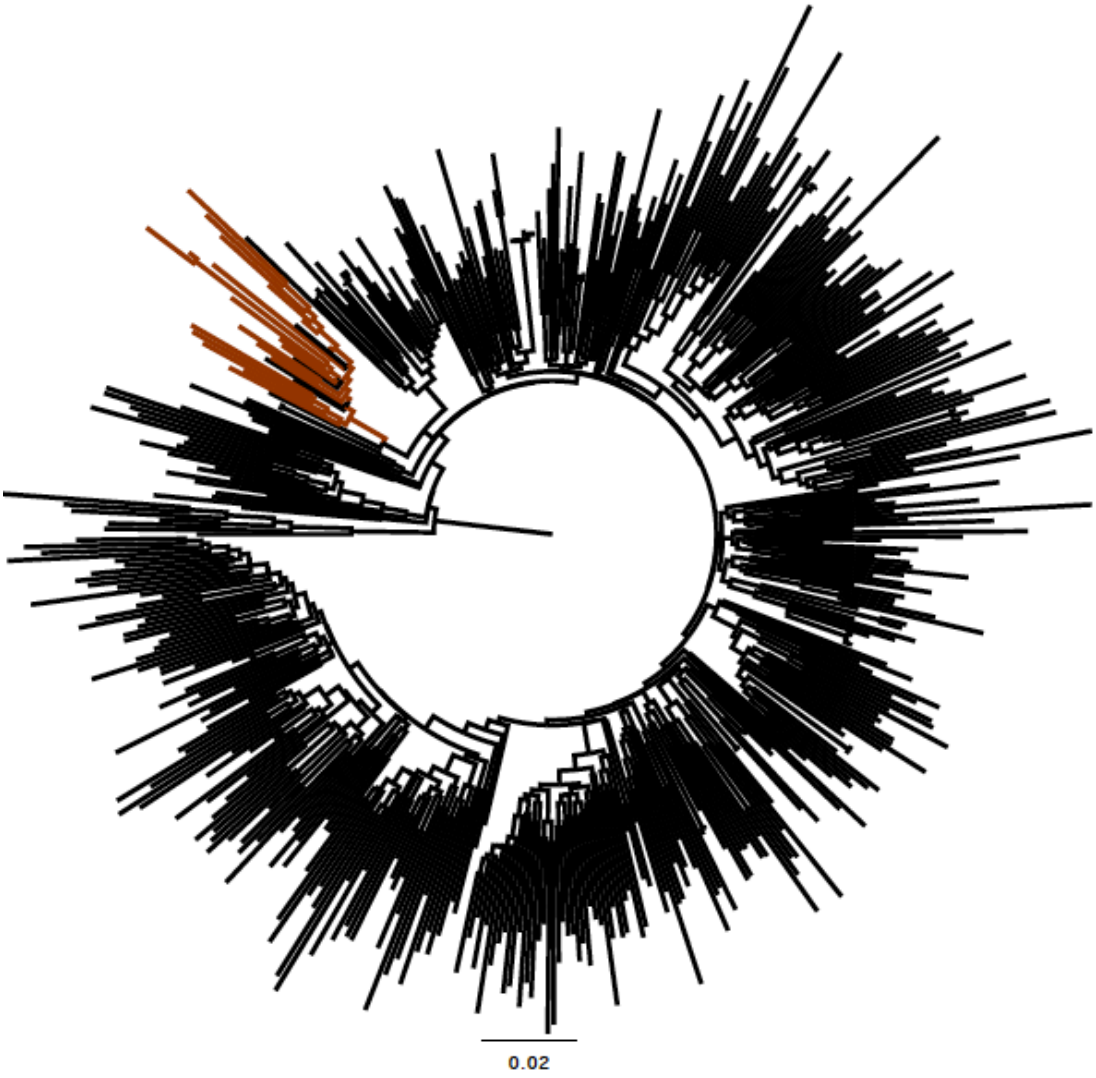


Figure 3.15 Maximum Likelihood tree using HIV-1 subtype C *pol* sequences (n = 515). The phylogeny is rooted with an HIV-1 subtype B isolate. Sequences from the Gambia (highlighted in brown) form a monophyletic cluster, along with four sequences isolated from Senegal (bootstrap value at node 0.95). Branch lengths and scale represent nucleotide substitutions per site.

3.4 DISCUSSION

Information on the diversity of HIV-1 in the Gambia and Guinea-Bissau is currently lacking and the current study has attempted to address this gap. Documentation of

the ongoing HIV-1 epidemic and sequence data from West Africa is important for several reasons. In a region where HIV-1 diversity is higher than in many other parts of sub-Saharan Africa, such information is required to maintain accurate viral diagnostics and sensitive viral load assays. HIV-1 subtypes may differ biologically in areas such as viral fitness [243, 244] and coreceptor usage (e.g. likelihood of switch from R5 to X4 usage) [245, 246]. These may in turn translate into higher risk of disease progression in certain subtypes and recombinant viruses could also have certain advantages over their parent strains. Studies in East Africa, using both prevalent and incident infections, have shown a higher risk of progression to AIDS and AIDS-related death in subtype D (and inter-subtype recombinant) -infected individuals when compared to subtype A-infected patients [213, 247]; even following adjustment for baseline viral load [248]. A Senegalese study supports the notion that non-A subtype infections progress faster than subtype A infections [249], although disease progression in CRF02_AG infected individuals appear to be no worse compared to non-AG infections [250]. This is despite the rise of this circulating recombinant form (CRF) in West Africa and *in vitro* data suggesting enhanced viral fitness [243]. With the increasing availability of anti-retroviral therapy (ART) in West Africa, it is also important to consider potential differences between HIV-1 subtypes in drug resistance pathways and the ease with which resistance appears due to naturally occurring polymorphisms (e.g. the development of K65R in subtype C infections) [214, 215]. Such findings would clearly have implications for local ART regimes and choice of 2nd line drugs. Reassuringly, the outcomes on ART in non-B subtypes appear to be broadly similar to subtype B infections, with a suggestion that CRF02_AG and subtype A infected individuals may even show more favorable responses (although both these studies were retrospective and based on European cohorts) [251, 252]. Finally, local

sequence data are important in the design of potential immunogens for future prophylactic and therapeutic HIV-1 vaccines, although the greater diversity in West Africa makes this daunting task even more challenging in this sub-region. Mosaic vaccine strategies [216, 217] may overcome this barrier and documentation of new CRFs and accurate representation of global sequence diversity is essential for these strategies.

Although CRF02_AG infections did predominate (approx. 50%) in the Gambia, the diversity of circulating HIV-1 subtypes was found to be high. Particularly interesting was the discovery of a novel CRF, now named CRF49_cpx in the LAHDB, which after CRF02_AG was found to be the most frequent cause of HIV-1 infections in subjects seen at the MRC Gambia GUM clinic. Although this cohort primarily represents an urban, coastal population in the Gambia, the clinic draws patients from throughout the country and may be a reasonable representation of nationwide HIV-1 subtype distribution. The diversity between the three CRF49_cpx near full genome sequences may suggest that they are derived from a virus that recombined decades ago and as a great deal of evolution may have occurred since that time, many of the recombination breakpoints cannot be clearly defined. Surprisingly, no individuals infected with CRF49_cpx are found (to date) in neighbouring countries. When a phylogeny was reconstructed including all unclassified *pol* sequences from Africa (n = 147), along with several CRF49_cpx isolates, no sequences were found that were more closely related to CRF49_cpx than the Botswana isolate 98BW21.17 [242] (data not shown). No sequences from Guinea-Bissau (in LAHDB or those in the current study) or Mali have regions consistent with CRF49_cpx viruses. Of the 895 individuals with sequence data obtained in Senegal available in the LAHDB, one individual is reported as having a

subtype A/J recombinant *gag* (415bp) and subtype J envelope (532bp) (accession numbers AY249795 and AY249789 respectively). It is possible that this represents a CRF49_cpx infection. Although there may be unreported CRF49_cpx infections in these countries, in the context of a reasonable number of available sequence data from Senegal, Mali and Guinea-Bissau, it is clear that the prevalence is extremely low compared to the Gambia. Given the geographical proximity and shared ethnic groups between these countries, this finding is unexpected. The origin of CRF49_cpx is also not clear. Intriguingly, the most similar known sequence is isolate 98BW21.17 from Botswana [242]. There are no known close social or cultural links between Botswana and the Gambia. The high prevalence of subtype C infections in Botswana (99.2% of sequences in the LAHDB) makes it likely that recombination between a 98BW21.17-like isolate and a subtype C isolate occurred in Southern Africa and subsequently acted as the origin of Gambian CRF49_cpx infections following introduction of this recombinant into the Gambia.

The HIV-1 subtype C isolates in the Gambia are monophyletic and therefore, like CRF49_cpx, it is possible that a single introduction followed by spread within the country accounts for the presence of subtype C infections in the Gambia. Recent data from a cohort of men who have sex with men (MSM) in Senegal, where a high proportion (40%) of HIV-1 infections are due to subtype C, suggests efficient spread following a single introduction into this MSM population has also occurred [253]. It is possible, therefore, that subtype C and CRF49_cpx infections in the Gambia represent spread within certain defined populations or risk groups to explain their restricted distribution. However, no clear associations between any demographic factors (including ethnicity) were found to explain these findings in the Gambia (data not shown), although detailed social and epidemiological data

are not available. Although as shown above, some subtype C isolates from Senegal did cluster with the Gambian subtype C sequences, these were distinct from the Senegalese MSM cluster (data not shown).

In contrast, multiple exchanges of CRF02_AG between the Gambia and other countries in West and Central Africa are evident. It appears that several introductions have contributed to CRF02_AG becoming the most prevalent HIV-1 subtype in the Gambia. Both previous studies [254] and data presented in Chapter 4 suggest that rapid dispersal of CRF02_AG occurred throughout West Africa soon after its emergence in the DRC. This early establishment of CRF02_AG in several countries in the sub-region would have, in turn, aided the continuing rise of CRF02_AG infections in West Africa, as regional migration and exchanges of HIV-1 infections took place. Nevertheless, the current study reveals how HIV-1 subtype distribution may differ significantly even in neighbouring countries. The proportion of infections caused by CRF02_AG is significantly higher and overall HIV-1 diversity lower in Guinea-Bissau when compared to the Gambia. This could be due to cultural differences, or to differences in the age and extent of the epidemic in each country. Of note, both trade and tourist links are more extensive in the Gambia than in Guinea Bissau and this is reflected by the observation that the net migration rates into the Gambia have been consistently one of the highest in the sub-region since 1950 (when compared with Guinea-Bissau, Guinea, Mali and Senegal [211]). An additional consideration was raised by the recent analysis concluding that CRF02_AG is more likely to be a pure subtype and the precursor to subtype G, which may in turn be a recombinant derived from subtypes CRF02_AG and J [255]. This history could account for the high prevalence of CRF02_AG in West Africa and may account for local differences (for example between Guinea-

Bissau and the Gambia) in the prevalence of subtype G and J viruses. A more recent analysis has however questioned these claims and suggested that CRF02_AG did indeed arise as a result of recombination events that occurred early in the divergence between subtype A and G [256]. Using the available sequence data in LAHDB for Senegal and Mali, the proportion of HIV-1 infections due to CRF02_AG in these countries is also higher than in the Gambia (63% and 70% respectively). It is important, however, to take into consideration the fact that the data available in the LAHDB is not within any strict epidemiological framework and this may significantly bias the HIV-1 subtype distribution observed in a particular country. In addition, the Senegalese data are dominated by sequences derived from specific cohorts (e.g. MSM, CSW), where particular subtypes may dominate [53, 257], while the Gambia data (mostly derived from the current study) come from random selection of patients attending a GUM clinic. Such differences in the patient composition could result in the differences observed in the subtype distribution. Nevertheless, it is clear that local sequence data are required to accurately document the circulating HIV-1 subtypes in each country, as distinct patterns may exist, such as the high proportion of CRF49_cpx in the Gambia and CRF06_cpx in Mali (approx. 11% of Mali sequences in LAHDB). The current study is likely to provide an accurate reflection of circulating subtypes in the cohorts studied, as the primers used were designed carefully to work across all subtypes. In addition, further primers were designed to try and amplify initial PCR failures.

Phylogenetic analyses of HIV-1 isolates from Caió also revealed multiple introductions of CRF02_AG infections into this rural community, although spread within Caió was also evident. Although Caió itself is relatively isolated geographically, its inhabitants travel frequently to the capital city Bissau to seek

work and visit family, as well as to Senegal, the Gambia and European countries such as France and Portugal. In fact, 48% of individuals from Caió included in this study admitted to having traveled outside Guinea-Bissau. The migratory nature of CSWs in Caió is also likely to be relevant in this context, as many travel to work in Bissau, Senegal, the Gambia and the truckers' route through eastern Guinea-Bissau [190]. It is unlikely that HIV-1 transmission between countries in West Africa due to migratory CSW is unique to Caió, as an early study found approximately 50% of CSWs working in the Gambia originated from countries outside of the Gambia [258]. These findings exemplify the complexity of HIV-1 transmission in this region. A recent UNAIDS/World Bank collaborative report stresses the need to understand how HIV-1 is transmitted between and within high-risk groups in West Africa, emphasising the paucity of knowledge that currently exists [259]. The report concludes that this information is crucial to help design evidence-based and targeted public health policies to halt the rising HIV-1 incidence in some West African countries. The heterogeneity in epidemics between neighbouring countries is also highlighted. The current study has attempted to improve the existing HIV-1 sequence database from the Gambia and Guinea-Bissau. Ongoing documentation of how the HIV-1 epidemics evolve in these countries is important, especially with wider availability of ART in the future. Furthermore, the current study provides some initial insight into the widespread and complex nature of HIV-1 spread in West Africa. A larger investigation of HIV-1 transmission in the sub-region is warranted, combining both molecular epidemiology and socio-demographic studies, in order to help curb an HIV-1 epidemic that has been overtaking HIV-2 in recent years.

CHAPTER 4 – POPULATION DYNAMICS OF HIV-2 IN RURAL GUINEA-BISSAU: COMPARISON WITH HIV-1 AND ONGOING TRANSMISSION AT THE HEART OF THE EPIDEMIC.

4.1 INTRODUCTION

While HIV-1 has spread globally, HIV-2 is confined mainly to West Africa, with reducing incidence and prevalence during the last two decades [48, 52]. Over a third of HIV-2 infected subjects from the community cohort in Caió, Guinea-Bissau were shown to have undetectable (<100 copies/ml) HIV-2 viral loads (VLs), with a mortality rate no higher than that of HIV-uninfected subjects over an 18-year period [75]. Although proviral loads are equivalent in HIV-1 and HIV-2 infection at similar disease stages, plasma VLs are approximately 30-fold lower in HIV-2-infected people [93, 94, 260]. Presumably as a direct consequence of lower VLs in plasma and genital secretions, both vertical (4% without antiretroviral therapy) and horizontal transmission of HIV-2 occurs less frequently [54-56].

Guinea-Bissau is considered to be the centre of the HIV-2 epidemic, with the highest recorded adult prevalence in the 1980s (8%), with up to 20% of those aged over 40 infected in the capital, Bissau [44, 261]. How such a poorly transmissible virus reached such high levels of prevalence is not fully understood, but unscreened blood transfusions, increased prostitution with displaced soldiers, along with spread via ritual excisions and mass vaccination campaigns during the War of Independence (1963 – 74) may have played a role [44, 57, 58, 262]. The prevalence of HIV-2 has steadily fallen, while the HIV-1 epidemic is now an increasing problem in West Africa [51]. Yet HIV-2 transmissions continue to occur

and the host or virus characteristics that determine ongoing HIV-2 spread on a background of reducing community prevalence are unknown. It is possible that one or more particularly virulent viral lineages may cause the bulk of new infections, as certain HIV-2 capsid motifs are significantly associated with higher VL [104].

The widespread availability of HIV-1 sequences and recent advances in molecular phylodynamics have allowed the characterisation of HIV-1 transmission networks in Western cohorts [263-265]. The lower numbers of HIV-2 infected individuals and the paucity of sequence data have precluded similar analyses of HIV-2, and no data exist on HIV-2 transmission in rural West Africa. In the current study, modern phylogenetic techniques have been applied to HIV-2 sequence and socio-demographic data from a well-characterised community cohort in rural Guinea-Bissau in an effort to answer these questions.

4.2 METHODS

Laboratory methods used are detailed in Chapter 2, sections 2.2 – 2.4 and 2.7 – 2.8.

Subject selection

Samples were obtained from the three Caió serosurveys, each covering approximately 75% of the adult population, that were carried out in 1989-1991, 1996-8 and 2006-2007 to estimate HIV prevalence and incidence [51, 266, 267]. Blood samples were collected and subjects were interviewed using standardised questionnaires addressing risk factors for infection. Incident infections were defined by the presence of an initially HIV-2 antibody negative sample and at least one subsequent HIV-2 antibody positive sample, although due to the long intervals

between serosurveys, a precise date of infection cannot be determined. Persons who were diagnosed with HIV-2 infection in the first serosurvey were defined as 'pre1989' cases and are likely to represent individuals infected during the early growth of the epidemic [44]. Information on sexual risk behaviour came from the most detailed questionnaire used in the 1996-8 survey [266]. None of the subjects in the current analysis were HIV-1/HIV-2 dually infected. All subjects were antiretroviral treatment naïve at the time of sampling. Serological and molecular algorithms used for diagnosis of HIV infection, and CD4 count/VL quantification have been described in Chapter 2, section 2.17.

Informed consent was obtained from all participants. Prior to 2003 documented verbal informed consent was obtained and from 2003 onwards, written informed consent was obtained. This was in line with the guidelines of the two oversight bodies that approved these studies (the Gambia Government/MRC Laboratories Joint Ethics Committee and the Ministry of Health of Guinea-Bissau).

Sequences

HIV-2 group A g*ag* (n = 86, 690bp, p26) and *env* (n = 70, 1350bp, C2 to gp41) sequences were available from a subset of participants sampled in 1996-8, 2003 or 2006-8 [104, 268]. As only 53 (51%) individuals had both *gag* and *env* available, the datasets were analysed independently. Additional HIV-2 *env* sequences (n = 68, 360bp) previously generated from 1991 samples [103, 269] were also used in some analyses. In addition, all non-Caió HIV-2 *env* sequences (covering the same region as the Caió *env* fragments) from West Africa with location and date of sampling available on GenBank were used to provide a global background of HIV-2 sequences (n = 41). Due to the paucity of available non-Caió HIV-2 sequences in

GenBank, a similar analysis was not possible for HIV-2 *gag*. HIV-1 CRF02_AG (the most prevalent subtype in Guinea-Bissau [211]) partial *env* sequences (830bp) from 56 Caió subjects sampled in 1996-7, 2003 or 2006-7 [268] were used to compare HIV-1/HIV-2 population dynamics in Caió. GenBank accession numbers of all sequences used are detailed in the appendix

Phylogenetic and phylodynamic analyses

Alignments of *gag* and *env* sequences were created using ClustalW2 [220, 221] and manually edited using Se-Al (Sequence Alignment Editor, v2.0a11) [222]. Reconstruction of phylogenies (*gag* and *env*) and changes in HIV-1 and HIV-2 effective population size (N_e) over time (*env*) were inferred by Bayesian MCMC inference, using BEAST v1.6.1 [270], under the GTR model of nucleotide substitution with gamma-distributed rate heterogeneity, a lognormal relaxed molecular clock and the flexible Bayesian Skyline coalescent model. Both strict and relaxed molecular clock models were tested independently and compared by computing a Bayes Factor (BF) [271]. The relaxed clock model provided a significantly better fit to the data, as indicated by a BF >20, and was retained for the analysis. Tips of the dated phylogenies were calibrated with the sampling year of the sequences. The N_e is the number of infections contributing to onward transmission at a particular time (rather than the number of actual infections).

For estimation of HIV-2 N_e, the alignment was complemented with 68 *env* sequences from 1991 and partitioned to allow the region spanned by these additional sequences to evolve under a separate set of model parameters from the rest of the alignment. A normal prior was set on the root height of the HIV-2 tree, with a mean of 67.86 years since the most recent tip (2007) and a standard

deviation of 16, based on the time of origin of the HIV-2 epidemic estimated by Lemey *et al*. [44].

The Bayesian MCMC searches were set to 65,000,000 (*gag* phylogeny), 100,000,000 (*env* phylogeny including 41 non-Caió sequences) or 50,000,000 iterations (Caió HIV-1 and HIV-2 *env* N_e estimations) with trees sampled every 1000^{th} generations. An MCCT was selected from the sampled posterior distribution with the program TreeAnnotator version 1.5.2 [228], after discarding trees corresponding to a 10% burn-in. Bayesian skyline plots were generated, after discarding a 10% burn-in, with the program Tracer v1.5 [229].

Caió clusters supported by a posterior probability of 1.0 were selected as significant transmission clusters for further analysis. Branch support alone was relied on for cluster identification, since no consensus exists for defining transmission clusters on the basis of HIV-2 *gag* and *env* genetic distance. Several individuals were sampled for the first time in later serosurveys and found to be HIV-2 infected (27% and 34% in *gag* and *env* datasets respectively). It is therefore impossible to know whether they were infected before or after 1989 and were excluded from analyses comparing incident and pre1989 infections.

4.3 RESULTS

HIV-2-infected Subjects

Sequences from 103 HIV-2 infected individuals were used for cluster identification. The characteristics of the study population are listed in Table 4.1. The majority (66%) was female and the mean age at HIV-2 diagnosis was 49 years, a reflection of the HIV-2 prevalence being highest in older women [48]. 39% were found to be

infected with HIV-2 in the earliest survey in 1989 and 29% had documented incident HIV-2 infection.

Table 4.1. Characteristics of 103 subjects included in the phylogenetic analysis of HIV-2 infection, Caió, Guinea-Bissau, 1989-2007. *Missing data for 7 subjects **Missing data for 9 subjects. The variables of living, age and marital status were taken from the time of diagnosis. Viral load and CD4% were taken from the most recent measurement (2003 or 2006-2008). *** Unless otherwise stated.

Characteristic	**N (%)*** **
Women	68 (66)
Resident in central area of village	77 (75)
Mean age at HIV-2 diagnosis in years (SD)	49 (14)
Ever married (at time of HIV-2 diagnosis)	84 (82)
Timing of infection	
Subjects known to be infected prior to 1989-91 survey	40 (39)
Subjects with documented incident infections after 1991	30 (29)
Subjects without documented infection in 1989 or incident infection	33 (32)
Median viral load in copies/ml (IQR)*	600 (50-7097)
Viral load categories	
Low :<100 copies/ml	31 (32)
Medium : 100 -10,000 copies/ml	38 (40)
High: >10,000 copies/ml	27 (28)
Mean CD4 percentage (SD)**	29 (12)
Subjects for whom *gag* sequence was available (%)	86 (83)
Subjects for whom *env* sequence was available (%)	70 (68)
Subjects for whom both *gag* and *env* sequence were available (%)	53 (51)

Changes in the HIV-2 and HIV-1 effective population sizes over time

Bayesian Skyline Plots were utilised to reconstruct both HIV-2 and HIV-1 population dynamics in Caió by estimating the N_e through time (Figure 4.1). No published data exist comparing the relative dynamics of HIV-1 and HIV-2 in the same community. For HIV-2 (Group A), the estimated time of the most recent

common ancestor (tMRCA) was 1947 (95% CI 1932; 1962), which is consistent with previous estimates [44], and the population experienced an exponential growth phase coinciding, as previously described [44], with the War of Independence (1963 – 1974), but also with the timing of mass smallpox vaccination campaigns and increased access to blood transfusions [14]. The circulating HIV-1 CRF02_AG strains sampled in Caió emerged at a much later date, with the tMRCA estimated at 1973 (95% CI 1956; 1986), and with a rapid growth phase during the same period when HIV-2 growth was stunted. The rate of evolution of HIV-2 *env* was 2.67×10^{-3} substitutions/site/year (95% CI 1.85×10^{-3}, 3.46×10^{-3}) compared with 4.1×10^{-3} substitutions/site/year (95% CI 2.39×10^{-3}, 5.81×10^{-3}) for HIV-1. Interestingly, the growth rates following introduction, as well as the N_e at the origin of the epidemic appear remarkably similar for both viruses. In addition, both HIV-1 and HIV-2 have reached a plateau at similar orders of magnitude.

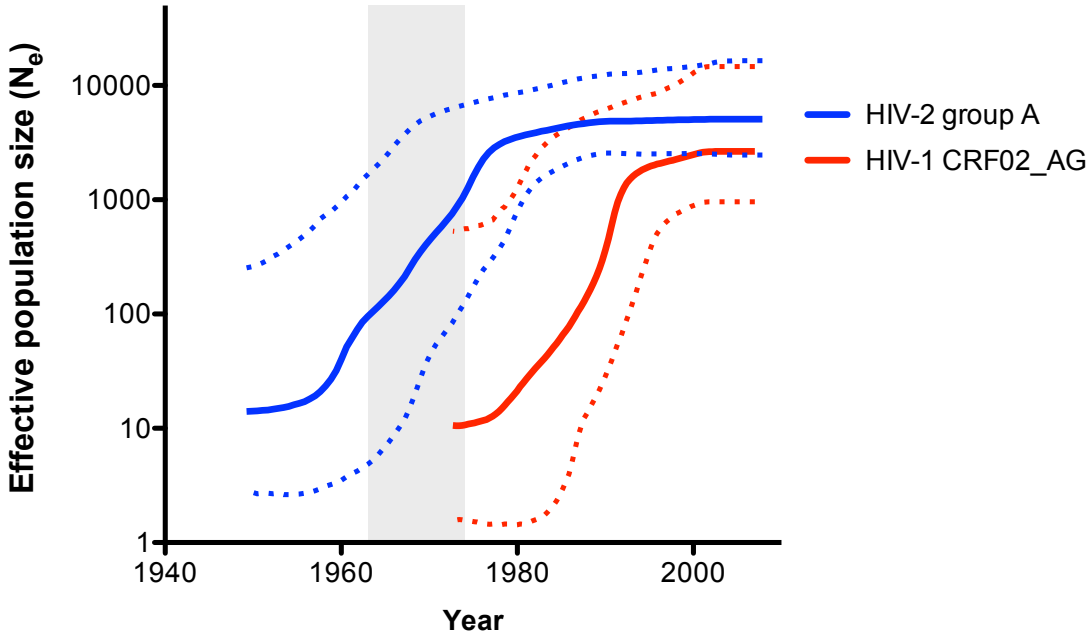


Figure 4.1. Bayesian Skyline Plots estimated for the HIV-2 group A and HIV-1 CRF02_AG epidemics in Caió, Guinea-Bissau. These represent estimates of the

effective population size (N_e) through time for HIV-2 and HIV-1 CRF02_AG *env* variants. Both plots commence at the mean posterior tMRCA, with the bold central lines representing the median N_e over time and dotted lines representing the 95% upper and lower confidence intervals. The shaded area represents the timeframe of the war of independence in Guinea-Bissau (1963 – 1974).

Identification of transmission clusters

From 86 *gag* sequences, 50 appeared in 13 transmission clusters – 1 cluster with 7 individuals, 4 with 5 individuals, 3 with 4 individuals, 1 with 3 individuals and 4 with 2 individuals. From 111 *env* sequences, 39/70 appeared in 13 (Caió-only) transmission clusters – 1 cluster with 6 individuals, 3 with 4 individuals, 3 with 3 individuals and 6 with 2 individuals (Table 4.2, Figure 4.2). This represents a high proportion of local individuals (58.1% in *gag* and 55.7% in *env*) involved in ongoing transmission of HIV-2 since the virus was introduced in Caió. Because the identified clusters span several decades and the datasets used represent only a proportion of the HIV-2 infected population in Caió, it is difficult to infer direct transmission with confidence, with the exception of known sexual partners. Several Caió *env* sequences were observed clustering with non-Caió sequences in 7 clusters (Figure 4.3), with Cape Verdean sequences involved in several clusters.

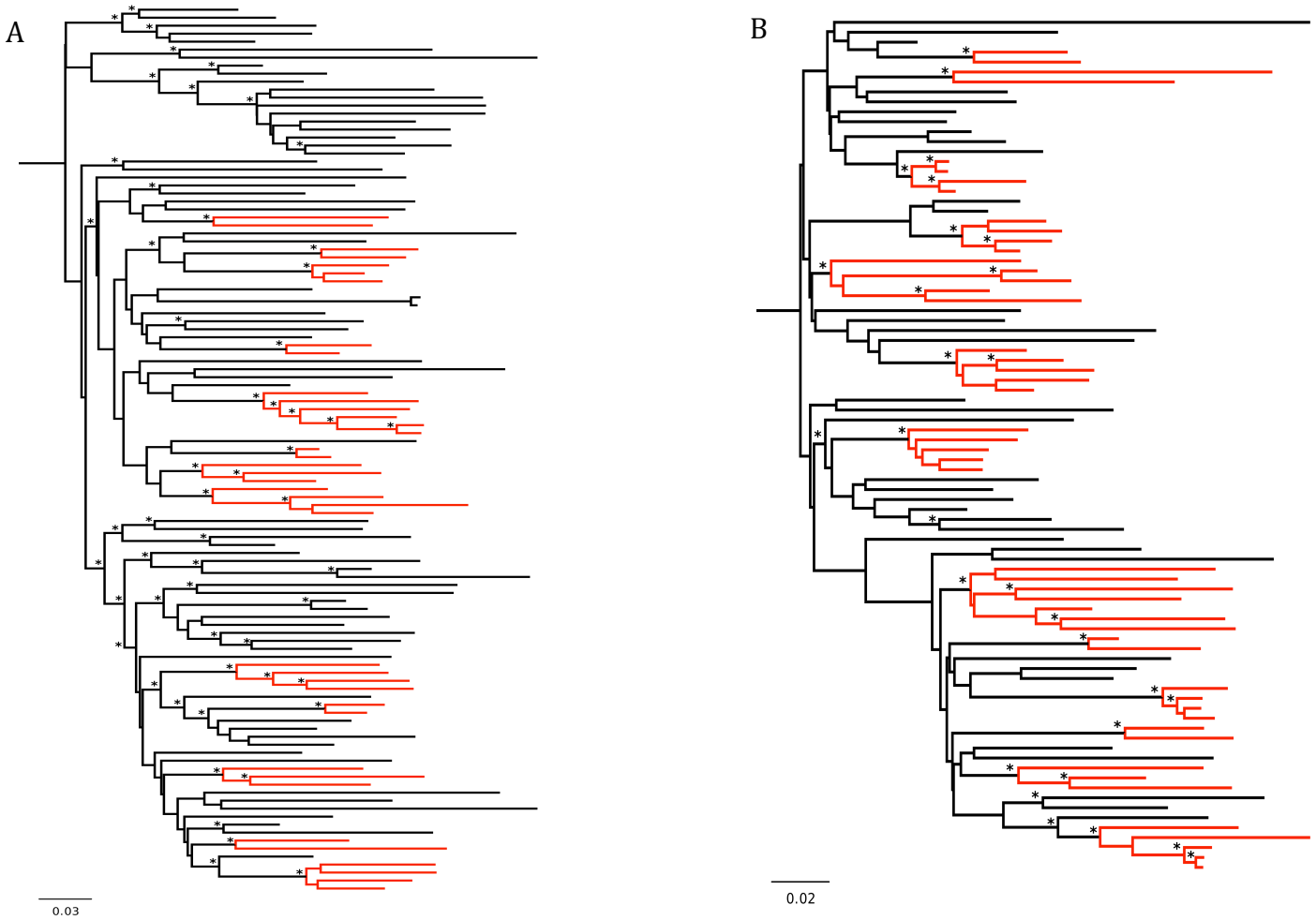


Figure 4.2. *env* (a) and *gag* (b) Maximum Clade Credibility Trees (MCCTs) of HIV-2 subtype A in Caió. Trees are midpoint rooted. Red lines indicate Caió transmission clusters (defined by a Bayesian posterior probability of 1.00). * indicates nodes with a posterior probability of >0.9. Scale bar indicates nucleotide substitutions per site.

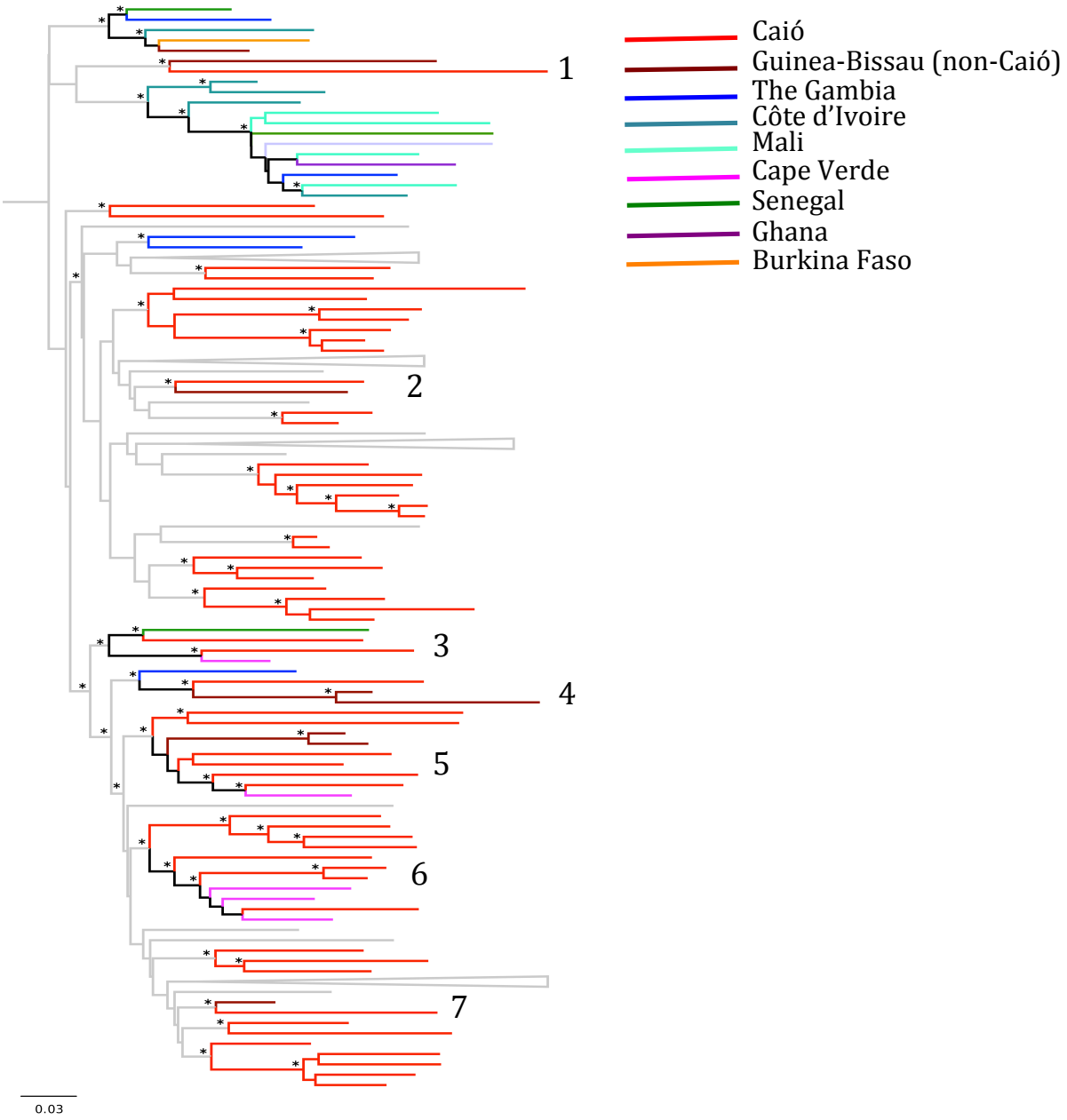


Figure 4.3. HIV-2 *env* Maximum Clade Credibility Tree (MCCT) phylogeny demonstrating clustering of Caió sequences with external HIV-2 isolates. The tree is midpoint rooted. In bold are clusters with a Bayesian posterior probability of ≥ 0.95 and * indicates nodes with a posterior probability of >0.9. Where possible, clades not supported by a posterior probability of >0.9 are collapsed. Scale bar

indicates nucleotide substitutions per site. Clusters with a mix of both Caió and non-Caió isolates are numbered.

Characteristics of clustering and non-clustering HIV-2 infections

No statistically significant differences in risk behaviour characteristics were found between clustering and non-clustering individuals (Table 4.2). For both phylogenies, clusters contained significantly more known sexual partners ($p<0.001$, Table 4.2). In the *gag* tree, individuals with clustering infections had significantly higher VLs than those with non-clustering infections ($p = 0.02$). Although a similar trend was observed in the *env* tree, this difference in median VLs was not significant. It is not clear to what extent infection with viruses with inherently different replicative capacity accounts for individuals with either undetectable HIV-2 plasma VL ('elite controllers') or progressive HIV-2 infection with detectable viraemia. In both *gag* and *env* phylogenies, elite controllers were found to share an MRCA with viraemic individuals within transmission clusters (in some cases these included known sexual partners, Figure 4.4). This is an important observation and it is likely, therefore, that the dichotomous outcomes in HIV-2 infection are largely determined by host factors.

Table 4.2. Characteristics of subjects in transmission clusters and those with non-clustering infections in Maximum Clade Credibility Trees (MCCTs) for *env* and *gag*. VL = Viral Load; CSW = Commercial Sex Worker; TPHA = Treponema Pallidum Haemagglutination test, HTLV-1 = Human T-Lymphotropic Virus-1, SD = standard deviation. *6 subjects with missing data for *env*; 2 subjects with missing data for *gag*. ** 5 subjects with missing data for *env*; 9 subjects with missing data for *gag.* Most recent data were used for VL and CD4% (2003 or 2006-8).

	Env-tree (n=70)			Gag-tree (n=86)		
	Cluster N=39	**Non-cluster N=31**	**p-value**	**Cluster N=50**	**Non-cluster N=36**	**p-value**
Demographics						
Women (%)	23 (59)	19 (61)	0.8	34 (68)	25 (69)	1.0
Mean age at diagnosis in years (SD)	48 (16)	49 (13)	0.8	48 (14)	47 (15)	0.6
Ever married (at time of diagnosis) (%)	29 (74)	24 (77)	0.3	41 (82)	32 (89)	0.5
Having known sexual relationship with other subjects in the study (%)	14 (42)	0 (0)	<0.001	16 (32)	2 (6)	0.003
Ever lived outside of Caió (%)	36 (92)	28 (90)	0.5	45 (90)	31 (86)	0.6
Sexual behaviour, STIs & other risk factors						
Ever worked as a CSW (of women) (%)	3 (13)	1 (5)	0.4	3 (9)	4 (17)	0.4
Ever visited a CSW (of men) (%)	8 (50)	4 (33)	0.1	4 (36)	6 (40)	0.6
Had ≥2 sexual partners in preceding year (%)	12 (31)	5 (16)	0.2	12 (24)	9 (25)	1.0
Ever had blood transfusion (%)	2 (5)	1 (3)	0.8	2 (6)	1 (2)	0.8
TPHA positive (%)	15 (38)	9 (29)	0.4	20 (40)	12 (33)	0.9
Ever had a genital sore (%)	9 (23)	5 (16)	0.1	9 (18)	7 (19)	1.0
HTLV-1 positive (%)	4 (10)	5 (16)	0.5	6 (12)	2 (6)	0.5
Diagnoses and HIV-2 disease parameters						
HIV-2 infected prior to 1989-1991 (%)	6 (15)	16 (52)	<0.001	13 (26)	23 (64)	<0.001
Incident HIV-2 after 1991 (%)	17 (44)	4 (13)		21 (42)	5 (14)	
Median recent VL (IQR)*	2653 (387-17067)	1343 (267-10752)	0.6	1347 (50-9819)	179 (50-732)	0.02
Mean recent CD4% (SD)**	26 (12)	30 (9)	0.1	27 (13)	32 (9)	0.1

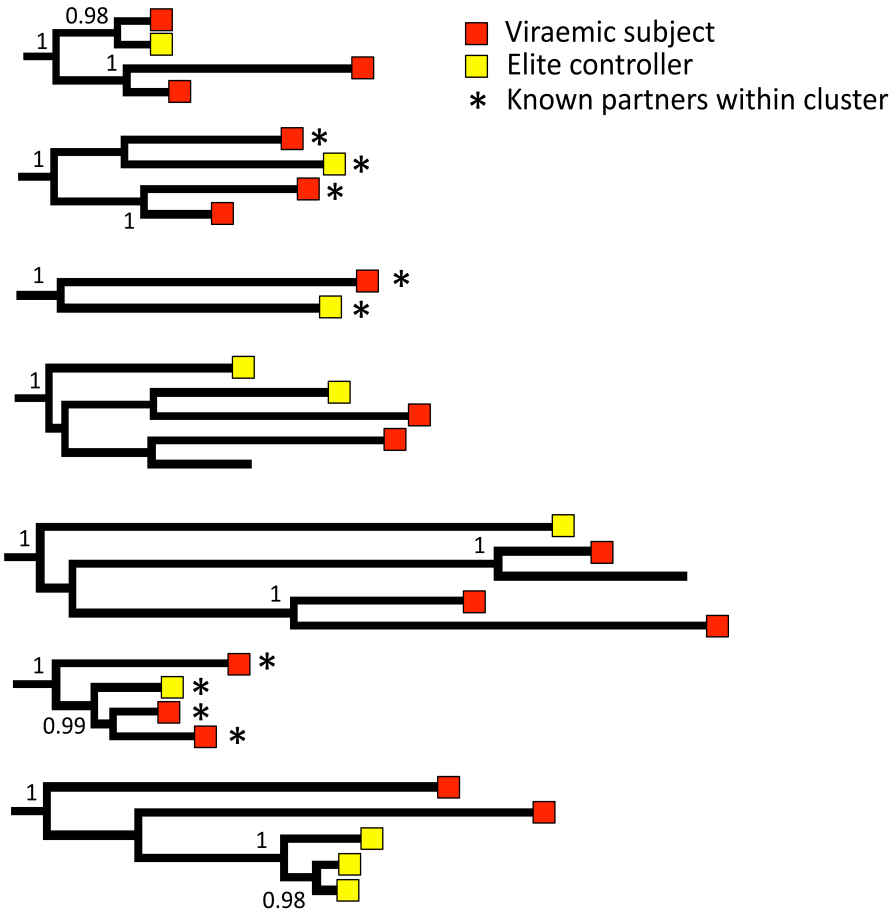


Figure 4.4. Selected transmission clusters from the *gag* phylogeny demonstrating discordant viral load status among individuals who share a most recent common ancestor (MRCA). Numbers displayed represent Bayesian posterior probabilities at each node. A similar pattern was found in clusters within the *env* phylogeny. Missing tip labels represent individuals where viral load data was not available.

Incident infections contribute significantly to HIV-2 transmission clusters

Transmission clusters contained significantly more incident infections than pre1989 cases (Table 4.2). A high proportion of all incident cases were also found

in clusters: 81.5% in *gag* and 83% in *env* phylogenies, compared to pre1989 cases: 36% in *gag* and 27% in *env* phylogenies; p = 0.0004 and p = 0.0003 respectively. Incident infections were not confined to one cluster, but appeared throughout the tree (Figure 4.5), suggesting there is not a single or few viral lineages that are responsible for ongoing HIV-2 transmissions. Interestingly, only one incident case significantly clustered with a non-Caió *env* isolate, whereas all others were found within Caió-only clusters. This suggests incident cases of HIV-2 in Caió are due to local transmission.

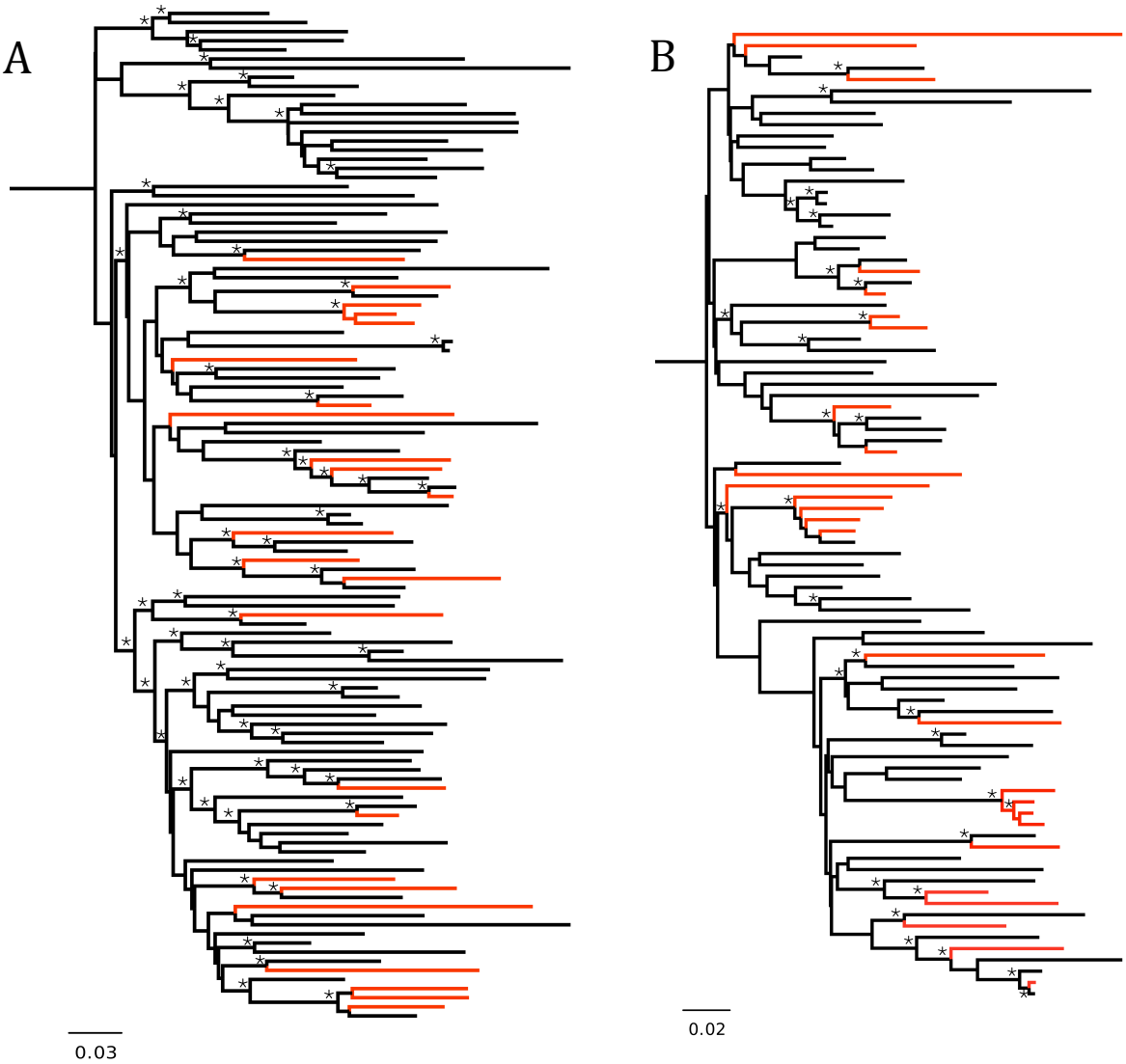


Figure 4.5. *env* (a) and *gag* (b) Maximum Clade Credibility Tree (MCCT) phylogenies of HIV-2 subtype A sequences. Trees are midpoint rooted. Branches in red indicate incident HIV-2 cases, which are distributed throughout the phylogeny. * indicates nodes with a posterior probability of >0.9. Scale bar indicates nucleotide substitutions per site.

While these data strongly suggest that after 1989, newly infected subjects are more strongly associated with ongoing HIV-2 transmission than older, prevalent infections in Caió, it is possible that sampling bias (i.e. exclusion of pre1989 individuals who had died by the time of subsequent sampling dates) may account for this result. A further *env* phylogeny was therefore reconstructed following inclusion of previously generated sequence data from 68 additional 1991 samples, which were all pre1989 infections [103, 269]. Although the use of these shorter sequences resulted in a phylogeny with poorer resolution, it is clear that the additional pre1989 cases formed clusters almost exclusively with other pre1989 cases, which would not be observed if the old infections were contributing to ongoing transmission in the cohort (Figure 4.6). This suggests that *ongoing* HIV-2 transmission in Caió occurs primarily from newly infected subjects rather than from the large pool of old infections in the community.

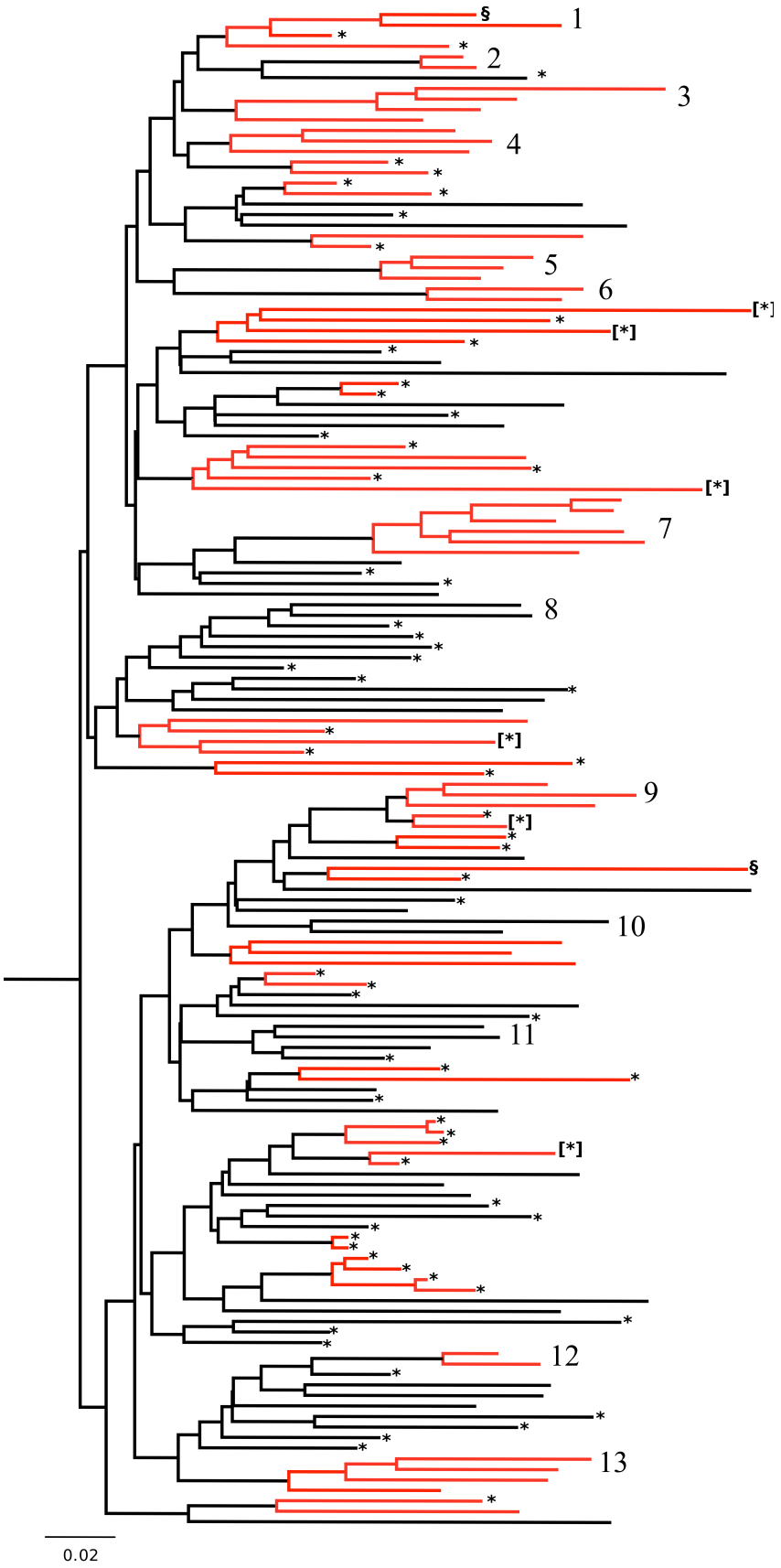


Figure 4.6. Maximum Clade Credibility Tree (MCCT) in nucleotide substitutions per site reconstructed using only Caió *env* sequences and

inclusion of additional individuals sampled in 1991. Significant clusters (posterior probability of 0.99 or higher at relevant nodes) are highlighted in red. The additional 68 sequences included from individuals samples in 1991 are denoted by a *. Caió Transmission clusters identified in the previous *env* phylogeny are denoted by numbers 1 – 13. The additional (n = 68) pre1989 sequences almost exclusively form new clusters with each other, or cluster with other pre1989 sequences included in the prior analysis (denoted by [*]; other than on two occasions where they form mixed clusters with incident cases (§).

4.4 DISCUSSION

The current study has investigated the phylodynamics and transmission of HIV-2 in rural Guinea-Bissau using sequence data from a well-characterised community cohort. Bayesian inference was used to analyse population dynamics of HIV-1 and HIV-2 in this cohort. The fact that the two epidemics experienced similar growth rates and patterns following introduction, but at different points in time, may suggest that the population dynamics are determined by high-risk host behaviour at onset rather than ecological or historical factors. Similar conclusions were drawn following comparison of the HIV-1 epidemics amongst men who have sex with men in the USA and UK [272]. Alternatively, the biological features that make HIV-2 less transmissible than HIV-1 may have been compensated for by factors such as a boost in transmission due to iatrogenic spread [57], allowing HIV-2 to reach epidemic growth rates akin to HIV-1. In the absence of this assistance, HIV-2 is declining due to inefficient heterosexual transmission. A third explanation could be that HIV-2 was more virulent and transmissible during the early phases of the epidemic, but due to poorer adaptability to hosts (than HIV-1), a gradual

attenuation has occurred over time. Data in Chapters 5 and 6, as well as previous reports from other groups demonstrate that HIV-2 is under significant evolutionary constraint when compared to HIV-1 [268, 273], which may account for HIV-2's disadvantageous competition with HIV-1. The current estimates of the HIV-1 CRF02_AG and HIV-2 group A tMRCA (1973 vs. 1947) are consistent with previous separate analyses using sequence data from Guinea-Bissau [44, 211]. The similarity between the tMRCA for Caió CRF02_AG isolates and the estimated emergence of CRF02_AG in the DRC emphasises the likelihood that rapid dispersal throughout West Africa occurred at an early stage [254].

The observation that the rapid rise of the HIV-1 N_e in this region coincided with plateauing of the HIV-2 N_e lends support to the theory that HIV-1 may have competitively displaced HIV-2; accounting for approximately 30% of the decrease in HIV-2 prevalence [274, 275]. The displacement of HIV-2 by HIV-1 on a population level may be attributed to increased mortality of a small susceptible pool of potential hosts due to HIV-1 infection. Additionally, HIV-1/HIV-2 dually infected individuals may transmit HIV-2 less readily since HIV-1 dominates the course of infection, resulting in very low HIV-2 VLs (reviewed in [276]). It is also possible that acquisition of HIV-1 protects against subsequent HIV-2 infection, due to factors such as cross-reactive immune responses, exhaustion of target cell population and CD4 down-regulation on HIV-1 infected cells. Although such a 'protective' effect has yet to be demonstrated, the sequence of infection in HIV-1/2 dually infected individuals is almost always HIV-2 followed by super-infection with HIV-1 (with no reported cases of HIV-1 infection followed by HIV-2). This of course may simply be due to the earlier presence of HIV-2 in the sub-region.

Establishing to what extent the near-disappearance of HIV-2 in West Africa is due to low VL, the presence of HIV-1 or other factors (e.g. risk behaviour, prevention campaigns) may have great relevance to future strategies in curbing the HIV-1 pandemic [277]. As recent interest has emerged in using early antiretroviral therapy to halt the spread of HIV-1 [278-280], some have suggested that with inherently low VLs, HIV-2 could be considered a natural model for what could be achieved in HIV-1 via early treatment on a large scale [277].

The current study also explored whether disease outcome (and plasma VL) in HIV-2 is associated with certain viral lineages. Two very distinct outcomes of HIV-2 infection have been observed. Approximately 30% of HIV-2 infected subjects control the infection and have undetectable VLs [75, 76], while others have high VLs, progress to AIDS and have mortality rates comparable to HIV-1 infected individuals [62, 70, 281, 282]. In the *gag* and *env* phylogenies, however, is was clear that individuals whose viruses share a MRCA can have discordant VL status (i.e. undetectable vs. detectable viraemia) suggesting that host factors (such as strong CD8+ T-cell responses [172] or genetic variation in innate immune responses) may play a greater role in determining elite control of HIV-2 and outcome of the infection. The observation that known sexual partners within transmission clusters sometimes have discordant VL status further strengthens this conclusion. A wider examination of known couples within the cohort revealed several other similarly discordant individuals (data not shown) and a comparative study of host immunity and genetics within these couples could provide valuable insight into what determines the outcome of HIV-2 infection.

The estimates of subject linkage in clusters (55 - 58%) are greater than the highest observed in the UK HIV-1 epidemic (36%) [283]. This is perhaps not surprising given this is a village cohort and case detection via community serosurveys is likely to identify a greater proportion of all those infected, and given that most of the identified clusters span several decades. Viruses from individuals infected since the start of the cohort in 1989 (*after* the exponential growth of HIV-2) are more frequently (81 - 83%) found in transmission clusters than those already infected by 1989 (27 - 36%). Although pre1989 individuals were involved in transmission events early in the epidemic, they have contributed less to more recent infections. In contrast, newer infections appear to be associated with ongoing HIV-2 transmission in Caió. It is also striking that even in a rural community in Guinea-Bissau, multiple lineages of HIV-2 are evident, with Caió isolates clustering with several external sequences. Notably, many are from Cape Verde, another country with Portuguese colonial links. Guinea-Bissau and Cape Verde formed a strategic alliance during the shared war of independence with Portugal and were united by political ties until 1980. A recent phylogeographic analysis has also highlighted the shared ancestry of these two countries during the early dispersal of HIV-2 [284]. The migratory nature of the Caió population in general, but particularly that of CSWs who often seek work in many locations in the sub-region [190], may also have contributed to this observation. Interestingly, as almost all new cases were involved in Caió-only clusters, this heterogeneity in HIV-2 viruses is likely to reflect early events that lead to the exponential increase in HIV-2.

The role of primary infection in transmitting HIV-1 is well documented both in Western [283, 285, 286] and African [287] settings, with up to 50% of viruses from primary HIV-1 infections being involved in transmission chains in a Canadian

cohort [285]. In a study from the UK, transmission during early infection appeared to be more important in the men who have sex with men (MSM) epidemic than in the heterosexual epidemic [263]. The increased risk of transmission during acute/early HIV-1 infection can be explained by the peak in viraemia during this period [288]. Little is known about primary HIV-2 infection and it has rarely been documented [78]. It is conceivable that a similar viraemic peak occurs prior to reaching a set point. As these set points are much lower in HIV-2 than in HIV-1 infection [79], including approximately 30% of infected people with plasma VL below 100 copies/ml [75], the main window for HIV-2 transmission may well be during early infection - a hypothesis that is supported by the data above. Individuals in clusters in the *gag* phylogeny had significantly higher VLs than non-clustering individuals, although only a non-significant trend was observed in the *env* phylogeny. The significance of this finding is also not clear, as the VLs used are likely to be from samples taken many years after transmission events. An association between clustering infections and higher viraemia has been shown in some HIV-1 cohorts [286], but not others [285]. Alternately, the higher proportion of new HIV-2 infections found in transmission clusters may reflect both acquisition and forward transmission of HIV-2 during a defined period of increased sexual risk-taking behaviour. If most HIV-2 transmissions do occur during acute/early infection, screening programs aimed at identifying and treating these cases could reduce the risk of onward transmission and help eliminate HIV-2 from West Africa altogether.

Heterosexual contact is the most frequent route of HIV-1 transmission in Africa and it is estimated that at least 60% of new infections occur within married or cohabiting couples [289]. In the current cohort, 33 - 42% of viruses in clusters

belonged to known heterosexual partners, suggesting that a large proportion of new HIV-2 infections also occur within marriages. This highlights the need for prevention specifically targeting known partners, especially if similar patterns are seen in HIV-1 transmission in Guinea-Bissau. In Caió, polygamous relationships are very common [290], as illustrated by one cluster consisting of viruses from one man and 3 wives (Figure 4.4). Although this study examined several other characteristics that may be associated with HIV-2 transmission, no association was found with increased sexual risk behaviour or co-infections. This could indicate these factors contribute less to the transmission risk than stage of HIV-2 infection or VL. However, the information available in the analysis was limited and derived from one time point only and under-reporting of sexual risk behaviour cannot be excluded.

The main limitation of this study is the relatively low number of individuals included, compared to country-wide studies such as the ones performed on UK cohorts [263]. Dense sampling of a population with high HIV-2 prevalence would offer an opportunity for a more in-depth HIV-2 transmission analysis, but is unfeasible as most HIV-2 cohorts in West Africa are rapidly shrinking [51, 52, 210]. A similar analysis of HIV-1 transmission in Caió would be of great interest. This could help design public health measures to halt an HIV-1 epidemic that is increasing and evolving in Guinea-Bissau [51, 291]. A recent UNAIDS/World Bank collaborative report has highlighted the need to understand HIV-1 transmission dynamics in West Africa, focusing on the likely complexity of transmission within and between risk groups [259].

In summary, this study in a rural West African community suggests that the population dynamics of HIV-2 and HIV-1 followed similar paths after introduction, although HIV-1 is now displacing HIV-2 as the dominant infection. Despite the poor transmissibility of HIV-2, new heterosexual transmissions still occur. Only time will tell whether these transmission events are sufficient to maintain a low-level epidemic in the region, or whether a gradual elimination will occur, of this sometimes lethal retrovirus, from Guinea-Bissau – a country that once had the highest recorded prevalence of HIV-2 infection in the world.

CHAPTER 5: ASSESSMENT OF AUTOLOGOUS AND HETEROLOGOUS NEUTRALISING ANTIBODY RESPONSES IN CHRONIC HIV-2 INFECTION AND ASSOCIATION WITH VIRAL CONTROL.

5.1 INTRODUCTION

Potent and broadly neutralising antibodies (Nabs) are crucial for defining an effective, sterilising HIV vaccine but induction of such Nabs has represented arguably the greatest challenge to investigators since the beginning of the HIV-1 epidemic. Significant advances have been made in understanding the Nab response to HIV-1 infection, including detailed description of its appearance in acute infection, rapid ongoing viral escape resulting in low contemporaneous autologous Nab titres (i.e. the failure to neutralise the currently circulating virus) despite the development of Nab breadth in some individuals and isolation of broadly neutralising human monoclonal antibodies (reviewed in [134]). Very few studies have explored the Nab response in HIV-2-infected individuals, and often these studies have been limited by small patient numbers and inclusion of HIV-2 progressors on ART, thus not allowing meaningful correlation with clinical status [141, 142]. A cross-sectional evaluation of nine HIV-2-infected individuals suggested that subjects with AIDS had lower autologous Nab titres than those with asymptomatic infection [141] and that titres may be higher than seen in HIV-1 infection. One of the only other descriptions of autologous Nab in four HIV-2-infected individuals indicated little change in longitudinal Nab titres, implying that *env* evolution and Nab escape may be limited [142]. A key obstacle in interpreting studies on HIV neutralisation has been the lack of a standardised and validated neutralisation assay, as well as difficulties in culturing primary viral isolates for

autologous Nab evaluation. The latter problem is even more pertinent in HIV-2 given the lower VLs and slower replication kinetics of some HIV-2 isolates [99, 100]. Use of molecularly cloned envelopes in single-cycle TZM-bl cell luciferase reporter gene assays have overcome these barriers and are now commonplace in HIV-1 Nab studies. Only one published study to date has applied these techniques in HIV-2 Nab assessment [145], comparing heterologous Nab titres in HIV-1 and asymptomatic HIV-2-infected Senegalese subjects and concluding that greater breadth but lower potency distinguish HIV-2 from HIV-1 infection. The lack of autologous Nab studies using cloned HIV-2 envelopes is likely to be due to the technical challenges in amplifying plasma RNA-derived envelopes from individuals with low or undetectable viraemia.

It is important to clarify, using these newer Nab assays, whether in fact more potent and broadly neutralising Nab responses are present in HIV-2 infection and to establish whether a potent autologous Nab response *is* a key distinguishing feature between HIV-2 controllers and those with higher viraemia. Equally important is documenting how HIV-2 Env biology and evolution may differ from that of the HIV-1 envelope in relation to the Nab response and could provide crucial insight into HIV pathogenesis.

5.2 METHODS

Laboratory methods used in this study are detailed in sections Chapter 2, sections 2.2 – 2.15.

Study population

Subjects with chronic HIV-2 infection and available plasma samples from the Caió community cohort were chosen for *env* amplification and Nab studies. Twenty-one subjects covering a broad range of HIV-2 disease outcomes were chosen for autologous neutralisation studies. All subjects had been infected for at least six years, were free of HIV-1 co-infection and were ART naïve at the time of sampling.

HIV-2 and HIV-1 env amplification to compare sequence diversity and selective pressure

Partial HIV-2 envelope sequences (C2 to gp41) from 76 Caió subjects were available as part of concurrent molecular epidemiological studies in the cohort (see Chapters 3 and 4). These included the 21 HIV-2 infected subjects used for autologous neutralisation studies. Partial HIV-1 envelope sequences (C2 to gp41) from 56 subjects infected with CRF02_AG were chosen from available Caió sequences for comparison. Sequences from 34 HIV-2 infected subjects have been published previously [104] (GenBank accession numbers GQ485517 – GQ485550) and the remaining sequences (a total of 111 HIV-2 and 56 HIV-1 *env*) were deposited under accession numbers JN863732 – JN863898.

HIV-2 human monoclonal antibodies

Four anti-HIV-2 human monoclonal antibodies (Mab), isolated from two Gambian HIV-2 sero-positive donors and the prototypic anti-HIV-2 envelope human Mab 1.7a (all kindly provided by James Robinson) were used in the study to further characterise the breadth and potency of the anti-HIV-2 humoral response against both laboratory adapted and primary HIV-2 envelopes. Although the exact epitope specificities of these Mabs are currently being determined, preliminary data suggest these antibodies target diverse regions of the HIV-2 envelope: 6.10B (CD4-

binding site), 6.10F (V3), 1.4H (CD4-induced), 1.4B (overlaps V3) and 1.7A (V4) ([292] and James Robinson, personal communication, MS in preparation).

Sequence analysis and tests for codon selection

Sequences were aligned using CLUSTALW2 [220, 221] and all alignments inspected and manually edited using Se-Al (Sequence Alignment Editor, v2.0a11) [222]. Ambiguous regions were deleted prior to further analyses. Pairwise genetic distances were calculated using PAUP* [223] using the best-fit model of substitution for the relevant alignments as selected by jModelTest [293]: HKY+I+G model of evolution for intra-patient HIV-2 gp140 alignments and GTR+I+G for both inter-patient partial C2 – gp41 HIV-1 and HIV-2 sets. HIV-1 and HIV-2 subtype was determined using the REGA HIV subtyping algorithm [232] and for HIV-1 confirmed with *pol* amplification where ambiguous results were obtained (See Chapter 2). Detection of N-linked glycosylation sites was performed using N-GlycoSite [294]. Diversity at each amino acid position in protein alignments was calculated using Shannon's entropy [295]. A maximum likelihood phylogenetic tree of all subject HIV-2 gp140 sequences was constructed using PAUP* [223] under the GTR+I+G model of nucleotide substitution. The statistical robustness of the ML topology was assessed by bootstrapping with 1000 replicates using the neighbour-joining method.

Sites under positive and negative selection in HIV-1 and HIV-2 *env* alignments were identified by comparison of synonymous and non-synonymous substitution rates using three different methods in the Datamonkey web-server [296]: single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL) and Random effects likelihood (REL) [297, 298]. Estimations were carried out using codon

models of evolution selected by a genetic algorithm [299] on the Datamonkey suite and all alignments tested for recombination using the GARD tool [300]. A single recombination breakpoint in each of the HIV-1 and HIV-2 partial C2 – gp41 *env* alignments and two breakpoints in the HIV-2 gp140 set were identified and therefore multiple partition datasets used for all analyses of selection. A p-value cut-off of 0.05 (SLAC and FEL) or Bayes Factor of 50 (REL) was used to define significant positive or negative selection at a codon.

5.3 RESULTS

Subject characteristics

Heterologous Nab studies were conducted using plasma samples from 40 subjects with chronic HIV-2 infection, not known to be closely epidemiologically linked to subjects from whom Env clones were derived. The median age (range) of these subjects was 60 years (35 – 84) and 28/40 (70%) were female. All plasma samples were taken in 2006 and (28/40) 70% of the cohort had been diagnosed with HIV-2 for at least 9 years. Autologous Nab studies were conducted with 69 plasma RNA-derived envelopes isolated from 21 HIV-2 infected subjects (TD01-21). Subjects had a median age of 56 years (range 31 – 84) at the time of sampling and 11/21 (52%) were female. Median CD4 count and VL (range) were 508 cells/μl (100 – 1754) and 11,675 copies/ml (< 100 – 320,272) respectively; full characteristics are given in Table 5.1. The wide range in CD4 count and VL of these subjects reflect the variation in HIV-2 disease progression observed, including five LTNP HIV-2 with VL <100 copies/ml (TD01-05) who had been infected with HIV-2 for at least 9 - 14 years. One further functional HIV-2 envelope was isolated, but only used in heterologous Nab studies due to insufficient autologous plasma (CA7205.8).

Phylogenetic analysis confirmed that *env* isolates from each patient clustered together as expected (Figure 5.1). Seven individuals were included in both the autologous and heterologous Nab cohorts.

Table 5.1 Details of HIV-2 infected individuals from whom 69 *env* variants for neutralisation studies were derived.

	Age at sample	Sex	Year of HIV-2 diagnosis	Year of sample	CD4 count cells/µl (%)^	Plasma Viral Load (copies/ml)^	No of *env* variants isolated
TD01	60	M	1997	2006	570 (35)	<100	1
TD02	58	F	1989	2003	1086 (44)	<100	3
TD03	54	M	1989	2003	816 (40)	<100	2
TD04	45	F	1989	2003	1754 (45)	<100	4
TD05	39	F	1997	2006	535 (35)	<100	4
TD06	34	M	1997	2003	742 (32)	667	5
TD07	84	F	1989	2006	615 (28)	1343	1
TD08	59	M	1989	2006	770 (29)	2078	1
TD09	70	M	1989	2003	883 (30)	3482	4
TD10	31	M	1997	2006	200 (18)	6431	1
TD11*	56	M	1997	2003	1355 (32)	7260	3
	59			2006	800 (30)	14104	1
TD12	52	F	1989	2006	285 (28)	9979	1
TD13	73	F	1997	2003	480 (29)	13371	3
TD14	33	F	1989	2006	1075 (31)	23889	1
TD15	56	F	1997	2003	287 (11)	43396	5
TD16*	40	F	1997	2003	236 (12)	76493	5
	43			2006	100 (9)	148593	4
TD17	56	M	1997	2003	373 (20)	109221	3
TD18	40	F	1997	2006	290 (13)	146284	1
TD19*	61	F	1997	2003	121 (15)	223924	3
	64			2006	160 (8)	37503	3
TD20	43	M	1997	2003	164 (12)	313188	5
TD21	72	M	1996	2003	476 (33)	320272	5

* Subjects TD11, TD16 and TD19 have clones from two different time points.
^ CD4 count and viral load are from the same sample time point form which *env* variants were derived.

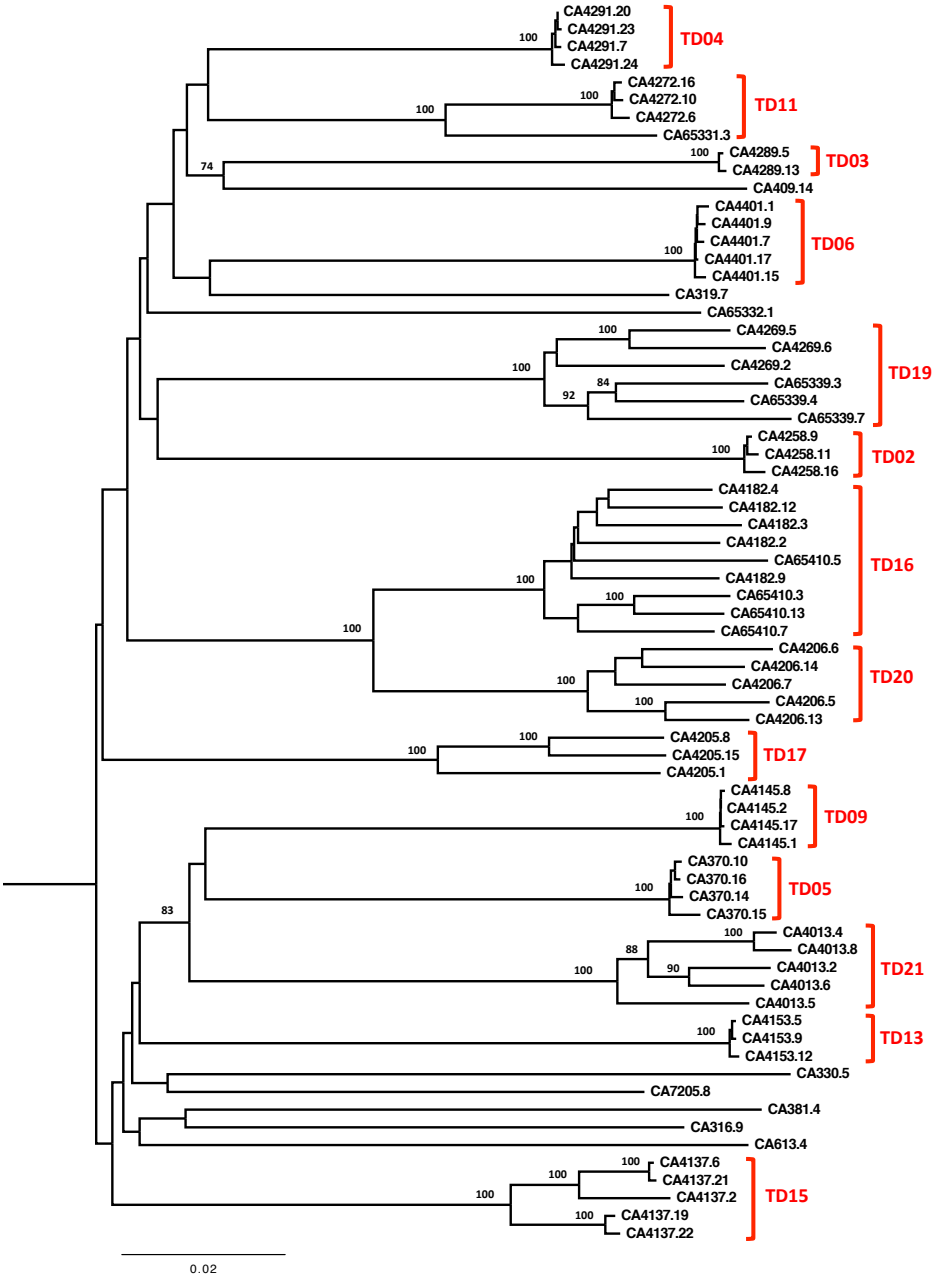


Figure 5.1 Midpoint rooted Maximum Likelihood tree demonstrating phylogenetic clustering of HIV-2 envelope clones from individual subjects from whom multiple clones were isolated. Bootstrap percentiles above 70% from 1000 replicates (using the Neighbour-Joining method) are shown at relevant branches. Branch lengths represent substitutions per nucleotide sites.

Broad and potent heterologous neutralising antibody responses are found in chronic HIV-2 infected subjects

To assess the breadth and potency of the heterologous Nab response in chronic HIV-2 infection, 40 HIV-2 plasma samples were tested against eight newly isolated primary HIV-2 Env variants (from eight different subjects) and two reference HIV-2 strains, 7312A (a kind gift from Beatrice Hahn) and ROD A, a known neutralisation-resistant laboratory-adapted HIV-2 isolate [143] (Figure 5.2). Six primary envelopes were gp160 cloned in pcDNA3.1/V5-His-TOPO® (CA330.5, CA316.9, CA319.7, CA381.4, CA7205.8, CA409.14) and two were gp140 clones in 7312A-SNAG (CA4205.8 and CA4182.4). Two envelopes (CA381.4 and CA409.14), representing extremes of neutralisation susceptibility phenotype, were compared in both gp140 and gp160 clone formats and as expected showed no appreciable difference in their sensitivity and neutralisation. Potent heterologous neutralising titres (IC_{50} above 1:7000) against 5 Caió Env variants and 7312A were seen with plasma from all (40/40) subjects, whereas ROD-A and three Caió Envs appeared to be relatively resistant to neutralisation (Figure 5.2).

Substantial breadth of heterologous neutralisation was also observed, with 35/40 subjects neutralising at least seven of the 10 Env variants tested and nine subjects able to neutralise eight or more Env variants. The magnitude of each subject's heterologous Nab response was defined as the median reciprocal IC_{50} titre against the 10 HIV-2 envelopes. Despite observing potent heterologous Nab in most subjects, a significant *positive* association (two-tailed, Spearman's rho = 0.359, p = 0.02) was seen between the magnitude of the heterologous Nab response and plasma VL against the panel of HIV-2 envelopes tested (Figure 5.3). A non-significant trend was observed towards higher heterologous Nab magnitude in

male compared to female subjects (median Nab IC_{50} 1:85557 vs 1:40890 respectively, p = 0.07). This may, however, be related to a similar trend observed for higher VLs in male subjects (median 2280 vs. 310 copies/ml, p = 0.05). This relationship between higher HIV-2 VL and male sex has been noted previously in the Caió cohort [76], but is currently unexplained. No gender difference in heterologous Nab breadth was observed.

Characteristics of clustering and non-clustering HIV-2 infections

No statistically significant differences in risk behaviour characteristics were found between clustering and non-clustering individuals (Table 4.2). For both phylogenies, clusters contained significantly more known sexual partners ($p<0.001$, Table 4.2). In the *gag* tree, individuals with clustering infections had significantly higher VLs than those with non-clustering infections ($p = 0.02$). Although a similar trend was observed in the *env* tree, this difference in median VLs was not significant. It is not clear to what extent infection with viruses with inherently different replicative capacity accounts for individuals with either undetectable HIV-2 plasma VL ('elite controllers') or progressive HIV-2 infection with detectable viraemia. In both *gag* and *env* phylogenies, elite controllers were found to share an MRCA with viraemic individuals within transmission clusters (in some cases these included known sexual partners, Figure 4.4). This is an important observation and it is likely, therefore, that the dichotomous outcomes in HIV-2 infection are largely determined by host factors.

Table 4.2. Characteristics of subjects in transmission clusters and those with non-clustering infections in Maximum Clade Credibility Trees (MCCTs) for *env* and *gag.* VL = Viral Load; CSW = Commercial Sex Worker; TPHA = Treponema Pallidum Haemagglutination test, HTLV-1 = Human T-Lymphotropic Virus-1, SD = standard deviation. *6 subjects with missing data for *env*; 2 subjects with missing data for *gag*. ** 5 subjects with missing data for *env*; 9 subjects with missing data for *gag.* Most recent data were used for VL and CD4% (2003 or 2006-8).

	Env-tree (n=70)			Gag-tree (n=86)		
	Cluster N=39	**Non-cluster N=31**	**p-value**	**Cluster N=50**	**Non-cluster N=36**	**p-value**
Demographics						
Women (%)	23 (59)	19 (61)	0.8	34 (68)	25 (69)	1.0
Mean age at diagnosis in years (SD)	48 (16)	49 (13)	0.8	48 (14)	47 (15)	0.6
Ever married (at time of diagnosis) (%)	29 (74)	24 (77)	0.3	41 (82)	32 (89)	0.5
Having known sexual relationship with other subjects in the study (%)	14 (42)	0 (0)	<0.001	16 (32)	2 (6)	0.003
Ever lived outside of Caió (%)	36 (92)	28 (90)	0.5	45 (90)	31 (86)	0.6
Sexual behaviour, STIs & other risk factors						
Ever worked as a CSW (of women) (%)	3 (13)	1 (5)	0.4	3 (9)	4 (17)	0.4
Ever visited a CSW (of men) (%)	8 (50)	4 (33)	0.1	4 (36)	6 (40)	0.6
Had ≥2 sexual partners in preceding year (%)	12 (31)	5 (16)	0.2	12 (24)	9 (25)	1.0
Ever had blood transfusion (%)	2 (5)	1 (3)	0.8	2 (6)	1 (2)	0.8
TPHA positive (%)	15 (38)	9 (29)	0.4	20 (40)	12 (33)	0.9
Ever had a genital sore (%)	9 (23)	5 (16)	0.1	9 (18)	7 (19)	1.0
HTLV-1 positive (%)	4 (10)	5 (16)	0.5	6 (12)	2 (6)	0.5
Diagnoses and HIV-2 disease parameters						
HIV-2 infected prior to 1989-1991 (%)	6 (15)	16 (52)	<0.001	13 (26)	23 (64)	<0.001
Incident HIV-2 after 1991 (%)	17 (44)	4 (13)		21 (42)	5 (14)	
Median recent VL (IQR)*	2653 (387-17067)	1343 (267-10752)	0.6	1347 (50-9819)	179 (50-732)	0.02
Mean recent CD4% (SD)**	26 (12)	30 (9)	0.1	27 (13)	32 (9)	0.1

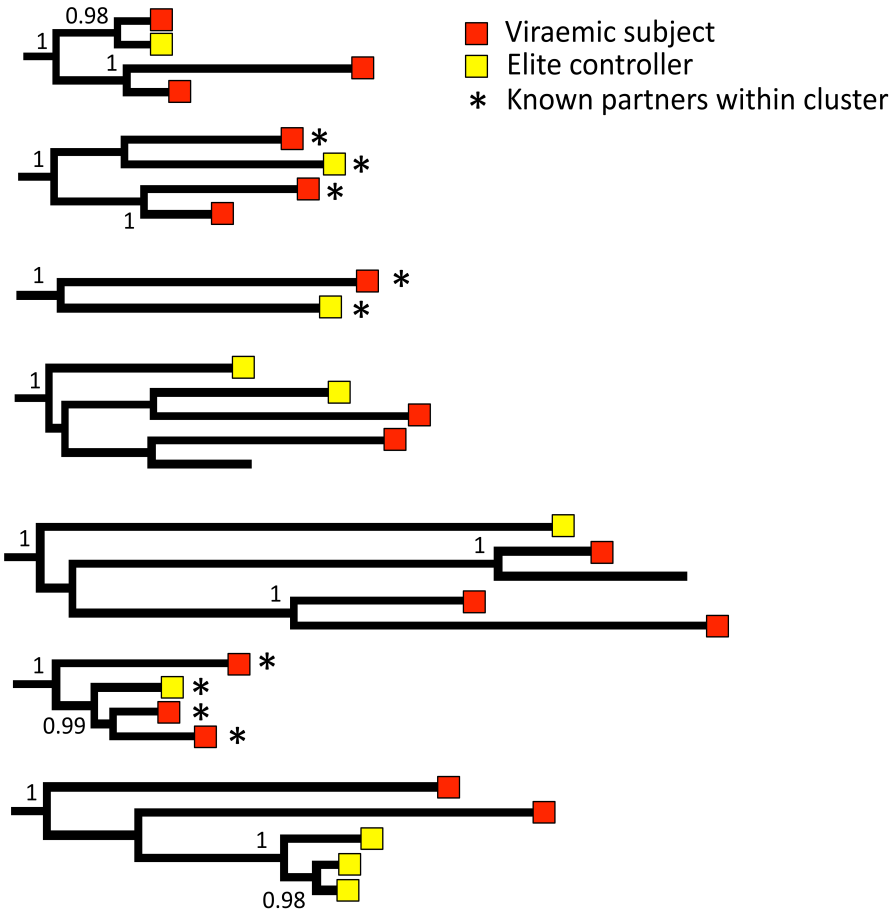


Figure 4.4. Selected transmission clusters from the *gag* phylogeny demonstrating discordant viral load status among individuals who share a most recent common ancestor (MRCA). Numbers displayed represent Bayesian posterior probabilities at each node. A similar pattern was found in clusters within the *env* phylogeny. Missing tip labels represent individuals where viral load data was not available.

Incident infections contribute significantly to HIV-2 transmission clusters

Transmission clusters contained significantly more incident infections than pre1989 cases (Table 4.2). A high proportion of all incident cases were also found

in clusters: 81.5% in *gag* and 83% in *env* phylogenies, compared to pre1989 cases: 36% in *gag* and 27% in *env* phylogenies; p = 0.0004 and p = 0.0003 respectively. Incident infections were not confined to one cluster, but appeared throughout the tree (Figure 4.5), suggesting there is not a single or few viral lineages that are responsible for ongoing HIV-2 transmissions. Interestingly, only one incident case significantly clustered with a non-Caió *env* isolate, whereas all others were found within Caió-only clusters. This suggests incident cases of HIV-2 in Caió are due to local transmission.

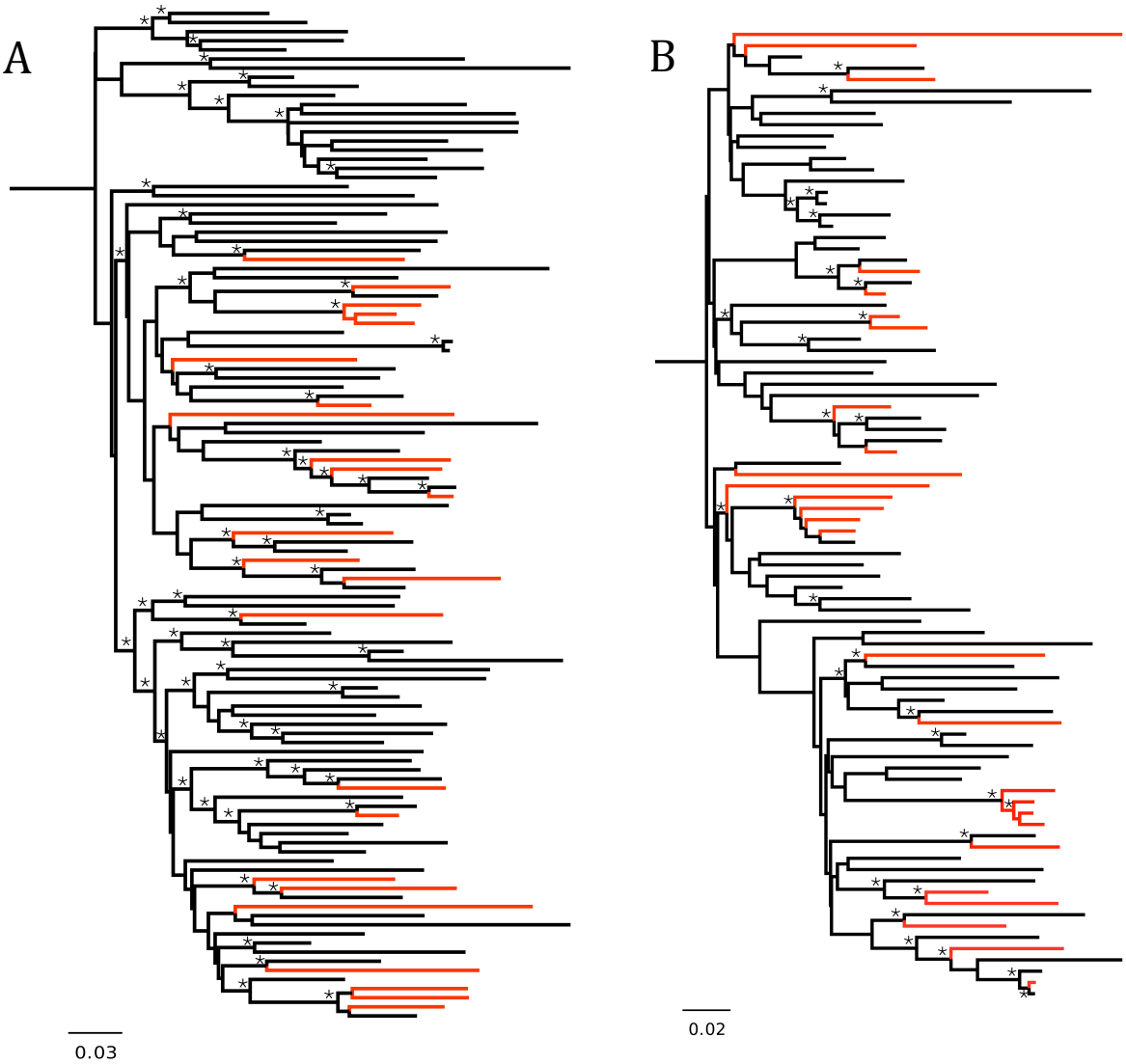


Figure 4.5. *env* (a) and *gag* (b) Maximum Clade Credibility Tree (MCCT) phylogenies of HIV-2 subtype A sequences. Trees are midpoint rooted. Branches in red indicate incident HIV-2 cases, which are distributed throughout the phylogeny. * indicates nodes with a posterior probability of >0.9. Scale bar indicates nucleotide substitutions per site.

While these data strongly suggest that after 1989, newly infected subjects are more strongly associated with ongoing HIV-2 transmission than older, prevalent infections in Caió, it is possible that sampling bias (i.e. exclusion of pre1989 individuals who had died by the time of subsequent sampling dates) may account for this result. A further *env* phylogeny was therefore reconstructed following inclusion of previously generated sequence data from 68 additional 1991 samples, which were all pre1989 infections [103, 269]. Although the use of these shorter sequences resulted in a phylogeny with poorer resolution, it is clear that the additional pre1989 cases formed clusters almost exclusively with other pre1989 cases, which would not be observed if the old infections were contributing to ongoing transmission in the cohort (Figure 4.6). This suggests that *ongoing* HIV-2 transmission in Caió occurs primarily from newly infected subjects rather than from the large pool of old infections in the community.

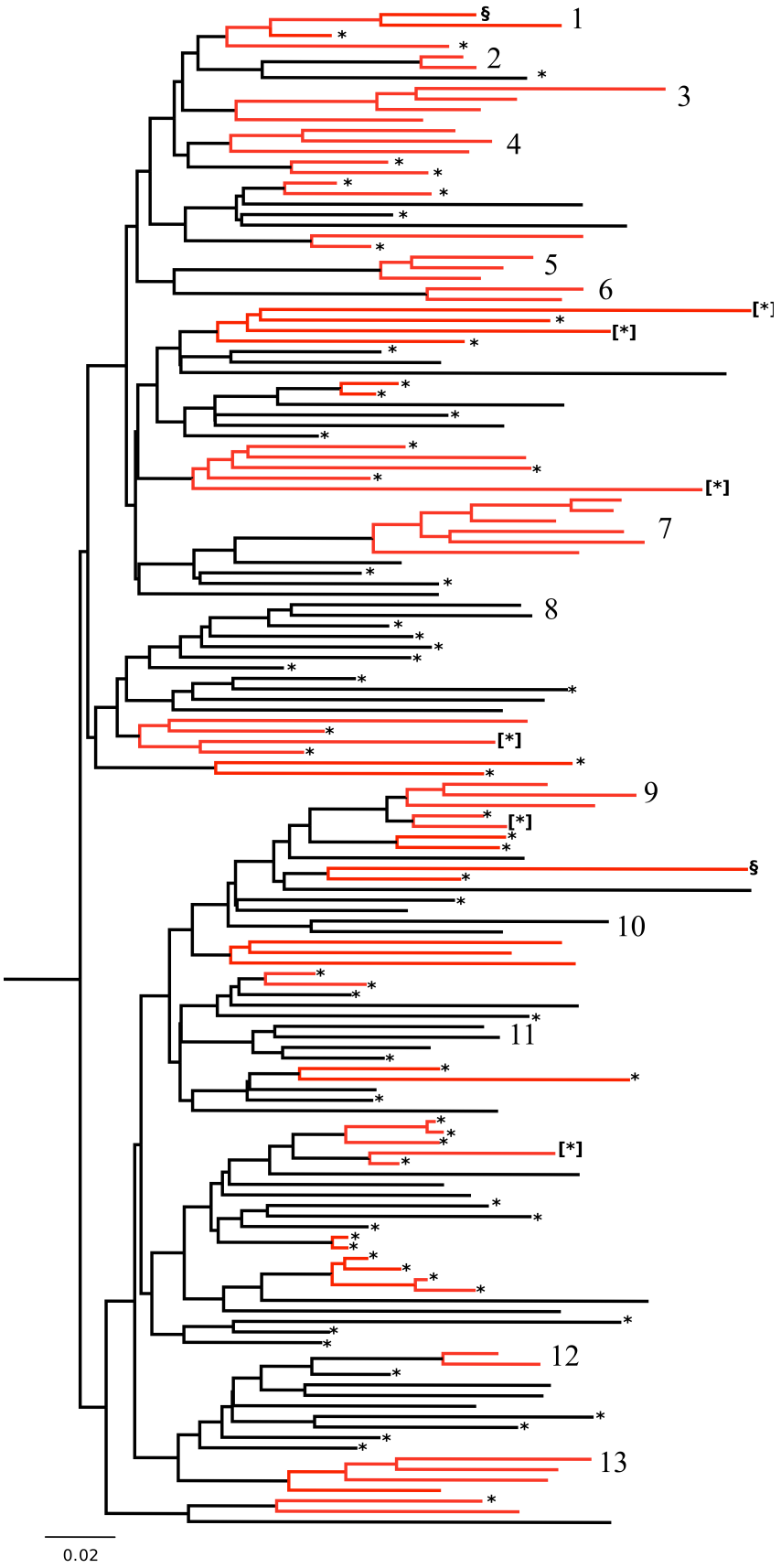


Figure 4.6. Maximum Clade Credibility Tree (MCCT) in nucleotide substitutions per site reconstructed using only Caió *env* sequences and

inclusion of additional individuals sampled in 1991. Significant clusters (posterior probability of 0.99 or higher at relevant nodes) are highlighted in red. The additional 68 sequences included from individuals samples in 1991 are denoted by a *. Caió Transmission clusters identified in the previous *env* phylogeny are denoted by numbers 1 – 13. The additional (n = 68) pre1989 sequences almost exclusively form new clusters with each other, or cluster with other pre1989 sequences included in the prior analysis (denoted by [*]; other than on two occasions where they form mixed clusters with incident cases (§).

4.4 DISCUSSION

The current study has investigated the phylodynamics and transmission of HIV-2 in rural Guinea-Bissau using sequence data from a well-characterised community cohort. Bayesian inference was used to analyse population dynamics of HIV-1 and HIV-2 in this cohort. The fact that the two epidemics experienced similar growth rates and patterns following introduction, but at different points in time, may suggest that the population dynamics are determined by high-risk host behaviour at onset rather than ecological or historical factors. Similar conclusions were drawn following comparison of the HIV-1 epidemics amongst men who have sex with men in the USA and UK [272]. Alternatively, the biological features that make HIV-2 less transmissible than HIV-1 may have been compensated for by factors such as a boost in transmission due to iatrogenic spread [57], allowing HIV-2 to reach epidemic growth rates akin to HIV-1. In the absence of this assistance, HIV-2 is declining due to inefficient heterosexual transmission. A third explanation could be that HIV-2 was more virulent and transmissible during the early phases of the epidemic, but due to poorer adaptability to hosts (than HIV-1), a gradual

attenuation has occurred over time. Data in Chapters 5 and 6, as well as previous reports from other groups demonstrate that HIV-2 is under significant evolutionary constraint when compared to HIV-1 [268, 273], which may account for HIV-2's disadvantageous competition with HIV-1. The current estimates of the HIV-1 CRF02_AG and HIV-2 group A tMRCA (1973 vs. 1947) are consistent with previous separate analyses using sequence data from Guinea-Bissau [44, 211]. The similarity between the tMRCA for Caió CRF02_AG isolates and the estimated emergence of CRF02_AG in the DRC emphasises the likelihood that rapid dispersal throughout West Africa occurred at an early stage [254].

The observation that the rapid rise of the HIV-1 N_e in this region coincided with plateauing of the HIV-2 N_e lends support to the theory that HIV-1 may have competitively displaced HIV-2; accounting for approximately 30% of the decrease in HIV-2 prevalence [274, 275]. The displacement of HIV-2 by HIV-1 on a population level may be attributed to increased mortality of a small susceptible pool of potential hosts due to HIV-1 infection. Additionally, HIV-1/HIV-2 dually infected individuals may transmit HIV-2 less readily since HIV-1 dominates the course of infection, resulting in very low HIV-2 VLs (reviewed in [276]). It is also possible that acquisition of HIV-1 protects against subsequent HIV-2 infection, due to factors such as cross-reactive immune responses, exhaustion of target cell population and CD4 down-regulation on HIV-1 infected cells. Although such a 'protective' effect has yet to be demonstrated, the sequence of infection in HIV-1/2 dually infected individuals is almost always HIV-2 followed by super-infection with HIV-1 (with no reported cases of HIV-1 infection followed by HIV-2). This of course may simply be due to the earlier presence of HIV-2 in the sub-region.

Establishing to what extent the near-disappearance of HIV-2 in West Africa is due to low VL, the presence of HIV-1 or other factors (e.g. risk behaviour, prevention campaigns) may have great relevance to future strategies in curbing the HIV-1 pandemic [277]. As recent interest has emerged in using early antiretroviral therapy to halt the spread of HIV-1 [278-280], some have suggested that with inherently low VLs, HIV-2 could be considered a natural model for what could be achieved in HIV-1 via early treatment on a large scale [277].

The current study also explored whether disease outcome (and plasma VL) in HIV-2 is associated with certain viral lineages. Two very distinct outcomes of HIV-2 infection have been observed. Approximately 30% of HIV-2 infected subjects control the infection and have undetectable VLs [75, 76], while others have high VLs, progress to AIDS and have mortality rates comparable to HIV-1 infected individuals [62, 70, 281, 282]. In the *gag* and *env* phylogenies, however, is was clear that individuals whose viruses share a MRCA can have discordant VL status (i.e. undetectable vs. detectable viraemia) suggesting that host factors (such as strong CD8+ T-cell responses [172] or genetic variation in innate immune responses) may play a greater role in determining elite control of HIV-2 and outcome of the infection. The observation that known sexual partners within transmission clusters sometimes have discordant VL status further strengthens this conclusion. A wider examination of known couples within the cohort revealed several other similarly discordant individuals (data not shown) and a comparative study of host immunity and genetics within these couples could provide valuable insight into what determines the outcome of HIV-2 infection.

The estimates of subject linkage in clusters (55 - 58%) are greater than the highest observed in the UK HIV-1 epidemic (36%) [283]. This is perhaps not surprising given this is a village cohort and case detection via community serosurveys is likely to identify a greater proportion of all those infected, and given that most of the identified clusters span several decades. Viruses from individuals infected since the start of the cohort in 1989 (*after* the exponential growth of HIV-2) are more frequently (81 - 83%) found in transmission clusters than those already infected by 1989 (27 - 36%). Although pre1989 individuals were involved in transmission events early in the epidemic, they have contributed less to more recent infections. In contrast, newer infections appear to be associated with ongoing HIV-2 transmission in Caió. It is also striking that even in a rural community in Guinea-Bissau, multiple lineages of HIV-2 are evident, with Caió isolates clustering with several external sequences. Notably, many are from Cape Verde, another country with Portuguese colonial links. Guinea-Bissau and Cape Verde formed a strategic alliance during the shared war of independence with Portugal and were united by political ties until 1980. A recent phylogeographic analysis has also highlighted the shared ancestry of these two countries during the early dispersal of HIV-2 [284]. The migratory nature of the Caió population in general, but particularly that of CSWs who often seek work in many locations in the sub-region [190], may also have contributed to this observation. Interestingly, as almost all new cases were involved in Caió-only clusters, this heterogeneity in HIV-2 viruses is likely to reflect early events that lead to the exponential increase in HIV-2.

The role of primary infection in transmitting HIV-1 is well documented both in Western [283, 285, 286] and African [287] settings, with up to 50% of viruses from primary HIV-1 infections being involved in transmission chains in a Canadian

cohort [285]. In a study from the UK, transmission during early infection appeared to be more important in the men who have sex with men (MSM) epidemic than in the heterosexual epidemic [263]. The increased risk of transmission during acute/early HIV-1 infection can be explained by the peak in viraemia during this period [288]. Little is known about primary HIV-2 infection and it has rarely been documented [78]. It is conceivable that a similar viraemic peak occurs prior to reaching a set point. As these set points are much lower in HIV-2 than in HIV-1 infection [79], including approximately 30% of infected people with plasma VL below 100 copies/ml [75], the main window for HIV-2 transmission may well be during early infection - a hypothesis that is supported by the data above. Individuals in clusters in the *gag* phylogeny had significantly higher VLs than non-clustering individuals, although only a non-significant trend was observed in the *env* phylogeny. The significance of this finding is also not clear, as the VLs used are likely to be from samples taken many years after transmission events. An association between clustering infections and higher viraemia has been shown in some HIV-1 cohorts [286], but not others [285]. Alternately, the higher proportion of new HIV-2 infections found in transmission clusters may reflect both acquisition and forward transmission of HIV-2 during a defined period of increased sexual risk-taking behaviour. If most HIV-2 transmissions do occur during acute/early infection, screening programs aimed at identifying and treating these cases could reduce the risk of onward transmission and help eliminate HIV-2 from West Africa altogether.

Heterosexual contact is the most frequent route of HIV-1 transmission in Africa and it is estimated that at least 60% of new infections occur within married or cohabiting couples [289]. In the current cohort, 33 - 42% of viruses in clusters

belonged to known heterosexual partners, suggesting that a large proportion of new HIV-2 infections also occur within marriages. This highlights the need for prevention specifically targeting known partners, especially if similar patterns are seen in HIV-1 transmission in Guinea-Bissau. In Caió, polygamous relationships are very common [290], as illustrated by one cluster consisting of viruses from one man and 3 wives (Figure 4.4). Although this study examined several other characteristics that may be associated with HIV-2 transmission, no association was found with increased sexual risk behaviour or co-infections. This could indicate these factors contribute less to the transmission risk than stage of HIV-2 infection or VL. However, the information available in the analysis was limited and derived from one time point only and under-reporting of sexual risk behaviour cannot be excluded.

The main limitation of this study is the relatively low number of individuals included, compared to country-wide studies such as the ones performed on UK cohorts [263]. Dense sampling of a population with high HIV-2 prevalence would offer an opportunity for a more in-depth HIV-2 transmission analysis, but is unfeasible as most HIV-2 cohorts in West Africa are rapidly shrinking [51, 52, 210]. A similar analysis of HIV-1 transmission in Caió would be of great interest. This could help design public health measures to halt an HIV-1 epidemic that is increasing and evolving in Guinea-Bissau [51, 291]. A recent UNAIDS/World Bank collaborative report has highlighted the need to understand HIV-1 transmission dynamics in West Africa, focusing on the likely complexity of transmission within and between risk groups [259].

In summary, this study in a rural West African community suggests that the population dynamics of HIV-2 and HIV-1 followed similar paths after introduction, although HIV-1 is now displacing HIV-2 as the dominant infection. Despite the poor transmissibility of HIV-2, new heterosexual transmissions still occur. Only time will tell whether these transmission events are sufficient to maintain a low-level epidemic in the region, or whether a gradual elimination will occur, of this sometimes lethal retrovirus, from Guinea-Bissau – a country that once had the highest recorded prevalence of HIV-2 infection in the world.

CHAPTER 5: ASSESSMENT OF AUTOLOGOUS AND HETEROLOGOUS NEUTRALISING ANTIBODY RESPONSES IN CHRONIC HIV-2 INFECTION AND ASSOCIATION WITH VIRAL CONTROL.

5.1 INTRODUCTION

Potent and broadly neutralising antibodies (Nabs) are crucial for defining an effective, sterilising HIV vaccine but induction of such Nabs has represented arguably the greatest challenge to investigators since the beginning of the HIV-1 epidemic. Significant advances have been made in understanding the Nab response to HIV-1 infection, including detailed description of its appearance in acute infection, rapid ongoing viral escape resulting in low contemporaneous autologous Nab titres (i.e. the failure to neutralise the currently circulating virus) despite the development of Nab breadth in some individuals and isolation of broadly neutralising human monoclonal antibodies (reviewed in [134]). Very few studies have explored the Nab response in HIV-2-infected individuals, and often these studies have been limited by small patient numbers and inclusion of HIV-2 progressors on ART, thus not allowing meaningful correlation with clinical status [141, 142]. A cross-sectional evaluation of nine HIV-2-infected individuals suggested that subjects with AIDS had lower autologous Nab titres than those with asymptomatic infection [141] and that titres may be higher than seen in HIV-1 infection. One of the only other descriptions of autologous Nab in four HIV-2-infected individuals indicated little change in longitudinal Nab titres, implying that *env* evolution and Nab escape may be limited [142]. A key obstacle in interpreting studies on HIV neutralisation has been the lack of a standardised and validated neutralisation assay, as well as difficulties in culturing primary viral isolates for

autologous Nab evaluation. The latter problem is even more pertinent in HIV-2 given the lower VLs and slower replication kinetics of some HIV-2 isolates [99, 100]. Use of molecularly cloned envelopes in single-cycle TZM-bl cell luciferase reporter gene assays have overcome these barriers and are now commonplace in HIV-1 Nab studies. Only one published study to date has applied these techniques in HIV-2 Nab assessment [145], comparing heterologous Nab titres in HIV-1 and asymptomatic HIV-2-infected Senegalese subjects and concluding that greater breadth but lower potency distinguish HIV-2 from HIV-1 infection. The lack of autologous Nab studies using cloned HIV-2 envelopes is likely to be due to the technical challenges in amplifying plasma RNA-derived envelopes from individuals with low or undetectable viraemia.

It is important to clarify, using these newer Nab assays, whether in fact more potent and broadly neutralising Nab responses are present in HIV-2 infection and to establish whether a potent autologous Nab response *is* a key distinguishing feature between HIV-2 controllers and those with higher viraemia. Equally important is documenting how HIV-2 Env biology and evolution may differ from that of the HIV-1 envelope in relation to the Nab response and could provide crucial insight into HIV pathogenesis.

5.2 METHODS

Laboratory methods used in this study are detailed in sections Chapter 2, sections 2.2 – 2.15.

Study population

Subjects with chronic HIV-2 infection and available plasma samples from the Caió community cohort were chosen for *env* amplification and Nab studies. Twenty-one subjects covering a broad range of HIV-2 disease outcomes were chosen for autologous neutralisation studies. All subjects had been infected for at least six years, were free of HIV-1 co-infection and were ART naïve at the time of sampling.

HIV-2 and HIV-1 env amplification to compare sequence diversity and selective pressure

Partial HIV-2 envelope sequences (C2 to gp41) from 76 Caió subjects were available as part of concurrent molecular epidemiological studies in the cohort (see Chapters 3 and 4). These included the 21 HIV-2 infected subjects used for autologous neutralisation studies. Partial HIV-1 envelope sequences (C2 to gp41) from 56 subjects infected with CRF02_AG were chosen from available Caió sequences for comparison. Sequences from 34 HIV-2 infected subjects have been published previously [104] (GenBank accession numbers GQ485517 – GQ485550) and the remaining sequences (a total of 111 HIV-2 and 56 HIV-1 *env*) were deposited under accession numbers JN863732 – JN863898.

HIV-2 human monoclonal antibodies

Four anti-HIV-2 human monoclonal antibodies (Mab), isolated from two Gambian HIV-2 sero-positive donors and the prototypic anti-HIV-2 envelope human Mab 1.7a (all kindly provided by James Robinson) were used in the study to further characterise the breadth and potency of the anti-HIV-2 humoral response against both laboratory adapted and primary HIV-2 envelopes. Although the exact epitope specificities of these Mabs are currently being determined, preliminary data suggest these antibodies target diverse regions of the HIV-2 envelope: 6.10B (CD4-

binding site), 6.10F (V3), 1.4H (CD4-induced), 1.4B (overlaps V3) and 1.7A (V4) ([292] and James Robinson, personal communication, MS in preparation).

Sequence analysis and tests for codon selection

Sequences were aligned using CLUSTALW2 [220, 221] and all alignments inspected and manually edited using Se-Al (Sequence Alignment Editor, v2.0a11) [222]. Ambiguous regions were deleted prior to further analyses. Pairwise genetic distances were calculated using PAUP* [223] using the best-fit model of substitution for the relevant alignments as selected by jModelTest [293]: HKY+I+G model of evolution for intra-patient HIV-2 gp140 alignments and GTR+I+G for both inter-patient partial C2 – gp41 HIV-1 and HIV-2 sets. HIV-1 and HIV-2 subtype was determined using the REGA HIV subtyping algorithm [232] and for HIV-1 confirmed with *pol* amplification where ambiguous results were obtained (See Chapter 2). Detection of N-linked glycosylation sites was performed using N-GlycoSite [294]. Diversity at each amino acid position in protein alignments was calculated using Shannon's entropy [295]. A maximum likelihood phylogenetic tree of all subject HIV-2 gp140 sequences was constructed using PAUP* [223] under the GTR+I+G model of nucleotide substitution. The statistical robustness of the ML topology was assessed by bootstrapping with 1000 replicates using the neighbour-joining method.

Sites under positive and negative selection in HIV-1 and HIV-2 *env* alignments were identified by comparison of synonymous and non-synonymous substitution rates using three different methods in the Datamonkey web-server [296]: single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL) and Random effects likelihood (REL) [297, 298]. Estimations were carried out using codon

models of evolution selected by a genetic algorithm [299] on the Datamonkey suite and all alignments tested for recombination using the GARD tool [300]. A single recombination breakpoint in each of the HIV-1 and HIV-2 partial C2 – gp41 *env* alignments and two breakpoints in the HIV-2 gp140 set were identified and therefore multiple partition datasets used for all analyses of selection. A p-value cut-off of 0.05 (SLAC and FEL) or Bayes Factor of 50 (REL) was used to define significant positive or negative selection at a codon.

5.3 RESULTS

Subject characteristics

Heterologous Nab studies were conducted using plasma samples from 40 subjects with chronic HIV-2 infection, not known to be closely epidemiologically linked to subjects from whom Env clones were derived. The median age (range) of these subjects was 60 years (35 – 84) and 28/40 (70%) were female. All plasma samples were taken in 2006 and (28/40) 70% of the cohort had been diagnosed with HIV-2 for at least 9 years. Autologous Nab studies were conducted with 69 plasma RNA-derived envelopes isolated from 21 HIV-2 infected subjects (TD01-21). Subjects had a median age of 56 years (range 31 – 84) at the time of sampling and 11/21 (52%) were female. Median CD4 count and VL (range) were 508 cells/μl (100 – 1754) and 11,675 copies/ml (< 100 – 320,272) respectively; full characteristics are given in Table 5.1. The wide range in CD4 count and VL of these subjects reflect the variation in HIV-2 disease progression observed, including five LTNP HIV-2 with VL <100 copies/ml (TD01-05) who had been infected with HIV-2 for at least 9 - 14 years. One further functional HIV-2 envelope was isolated, but only used in heterologous Nab studies due to insufficient autologous plasma (CA7205.8).

Phylogenetic analysis confirmed that *env* isolates from each patient clustered together as expected (Figure 5.1). Seven individuals were included in both the autologous and heterologous Nab cohorts.

Table 5.1 Details of HIV-2 infected individuals from whom 69 *env* variants for neutralisation studies were derived.

	Age at sample	Sex	Year of HIV-2 diagnosis	Year of sample	CD4 count cells/µl (%)^	Plasma Viral Load (copies/ml)^	No of *env* variants isolated
TD01	60	M	1997	2006	570 (35)	<100	1
TD02	58	F	1989	2003	1086 (44)	<100	3
TD03	54	M	1989	2003	816 (40)	<100	2
TD04	45	F	1989	2003	1754 (45)	<100	4
TD05	39	F	1997	2006	535 (35)	<100	4
TD06	34	M	1997	2003	742 (32)	667	5
TD07	84	F	1989	2006	615 (28)	1343	1
TD08	59	M	1989	2006	770 (29)	2078	1
TD09	70	M	1989	2003	883 (30)	3482	4
TD10	31	M	1997	2006	200 (18)	6431	1
TD11*	56	M	1997	2003	1355 (32)	7260	3
	59			2006	800 (30)	14104	1
TD12	52	F	1989	2006	285 (28)	9979	1
TD13	73	F	1997	2003	480 (29)	13371	3
TD14	33	F	1989	2006	1075 (31)	23889	1
TD15	56	F	1997	2003	287 (11)	43396	5
TD16*	40	F	1997	2003	236 (12)	76493	5
	43			2006	100 (9)	148593	4
TD17	56	M	1997	2003	373 (20)	109221	3
TD18	40	F	1997	2006	290 (13)	146284	1
TD19*	61	F	1997	2003	121 (15)	223924	3
	64			2006	160 (8)	37503	3
TD20	43	M	1997	2003	164 (12)	313188	5
TD21	72	M	1996	2003	476 (33)	320272	5

* Subjects TD11, TD16 and TD19 have clones from two different time points.
^ CD4 count and viral load are from the same sample time point form which *env* variants were derived.

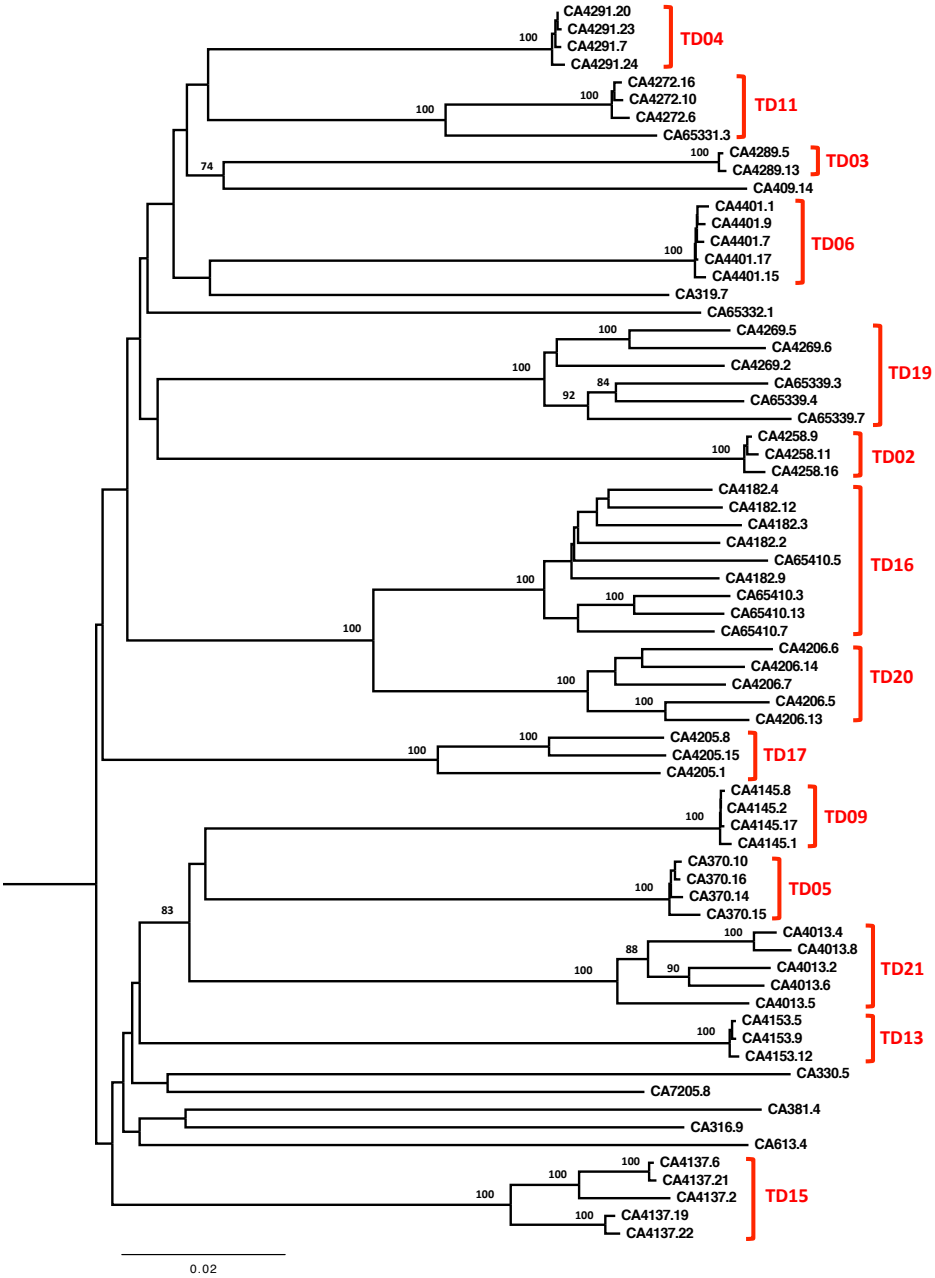


Figure 5.1 Midpoint rooted Maximum Likelihood tree demonstrating phylogenetic clustering of HIV-2 envelope clones from individual subjects from whom multiple clones were isolated. Bootstrap percentiles above 70% from 1000 replicates (using the Neighbour-Joining method) are shown at relevant branches. Branch lengths represent substitutions per nucleotide sites.

Broad and potent heterologous neutralising antibody responses are found in chronic HIV-2 infected subjects

To assess the breadth and potency of the heterologous Nab response in chronic HIV-2 infection, 40 HIV-2 plasma samples were tested against eight newly isolated primary HIV-2 Env variants (from eight different subjects) and two reference HIV-2 strains, 7312A (a kind gift from Beatrice Hahn) and ROD A, a known neutralisation-resistant laboratory-adapted HIV-2 isolate [143] (Figure 5.2). Six primary envelopes were gp160 cloned in pcDNA3.1/V5-His-TOPO® (CA330.5, CA316.9, CA319.7, CA381.4, CA7205.8, CA409.14) and two were gp140 clones in 7312A-SNAG (CA4205.8 and CA4182.4). Two envelopes (CA381.4 and CA409.14), representing extremes of neutralisation susceptibility phenotype, were compared in both gp140 and gp160 clone formats and as expected showed no appreciable difference in their sensitivity and neutralisation. Potent heterologous neutralising titres (IC_{50} above 1:7000) against 5 Caió Env variants and 7312A were seen with plasma from all (40/40) subjects, whereas ROD-A and three Caió Envs appeared to be relatively resistant to neutralisation (Figure 5.2).

Substantial breadth of heterologous neutralisation was also observed, with 35/40 subjects neutralising at least seven of the 10 Env variants tested and nine subjects able to neutralise eight or more Env variants. The magnitude of each subject's heterologous Nab response was defined as the median reciprocal IC_{50} titre against the 10 HIV-2 envelopes. Despite observing potent heterologous Nab in most subjects, a significant *positive* association (two-tailed, Spearman's rho = 0.359, p = 0.02) was seen between the magnitude of the heterologous Nab response and plasma VL against the panel of HIV-2 envelopes tested (Figure 5.3). A non-significant trend was observed towards higher heterologous Nab magnitude in

male compared to female subjects (median Nab IC_{50} 1:85557 vs 1:40890 respectively, p = 0.07). This may, however, be related to a similar trend observed for higher VLs in male subjects (median 2280 vs. 310 copies/ml, p = 0.05). This relationship between higher HIV-2 VL and male sex has been noted previously in the Caió cohort [76], but is currently unexplained. No gender difference in heterologous Nab breadth was observed.

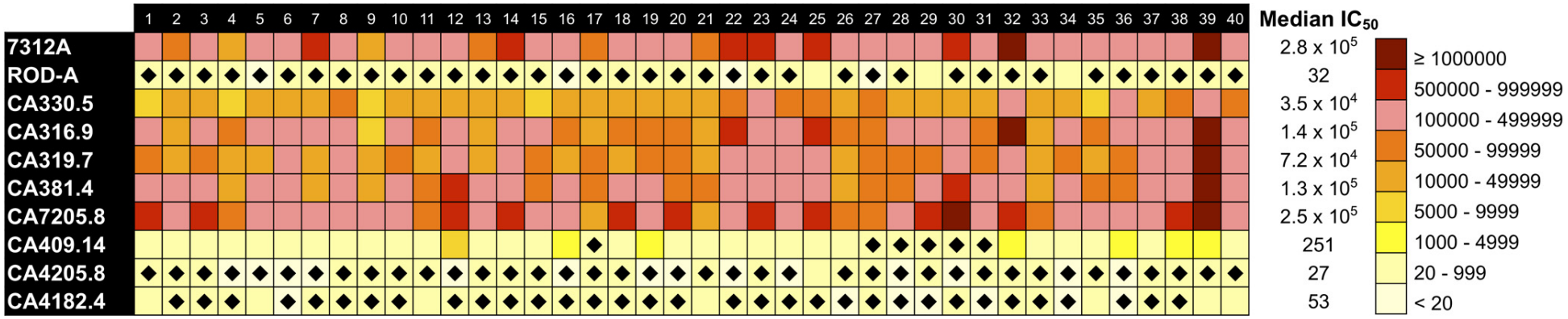


Figure 5.2 Heatmap depicting heterologous neutralisation titres of 40 HIV-2 infected individuals against 10 HIV-2 envelopes. Values given are reciprocal IC_{50}s. The panel clearly separates into envelopes with divergent neutralisation phenotypes. Plasma-envelope combinations that were resistant to neutralisation (based on a titre below a cutoff 3 times the VSV IC_{50}) are additionally marked with a closed diamond.

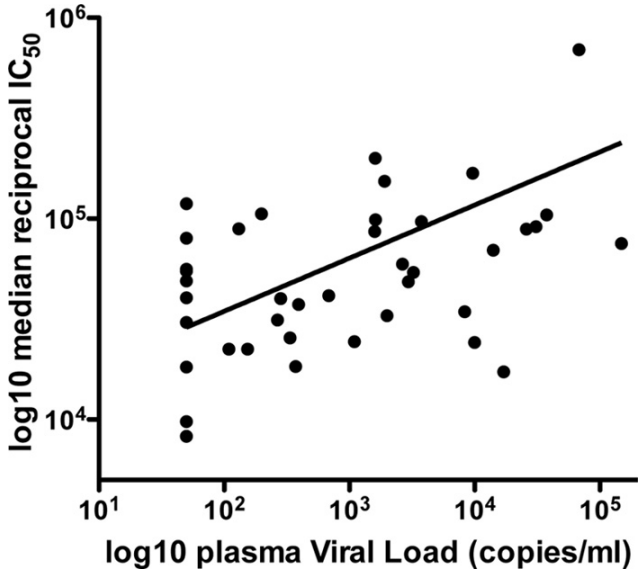


Figure 5.3 Relationship between plasma viral load and median reciprocal IC_{50} of heterologous neutralisation titres against a panel of 10 HIV-2 envelopes form HIV-2 infected individuals (n = 40). A significant positive correlation between viraemia and the magnitude of the heterologous neutralising antibody response is found (Spearman's rho = 0.359, p = 0.02).

Evidence for highly potent contemporaneous autologous neutralising antibody responses in chronic HIV-2 infection, but no role in viraemic control

Similar to the high magnitude heterologous Nab responses observed above, potent contemporaneous autologous Nab titres against 54/69 envelopes were observed, with IC_{50} titres above 1:10000 (Figure 5.4). However, no significant association was found between the magnitude of contemporaneous autologous Nab titres and plasma viraemia (Figure 5.4, Kruskal Wallis Test with Dunn's post-test for multiple comparisons, p = 0.0938). Notably, autologous Nab titres in HIV-2 controllers with

an undetectable plasma VL were equivalent to those seen in subjects with higher viraemia and HIV-2 disease progression.

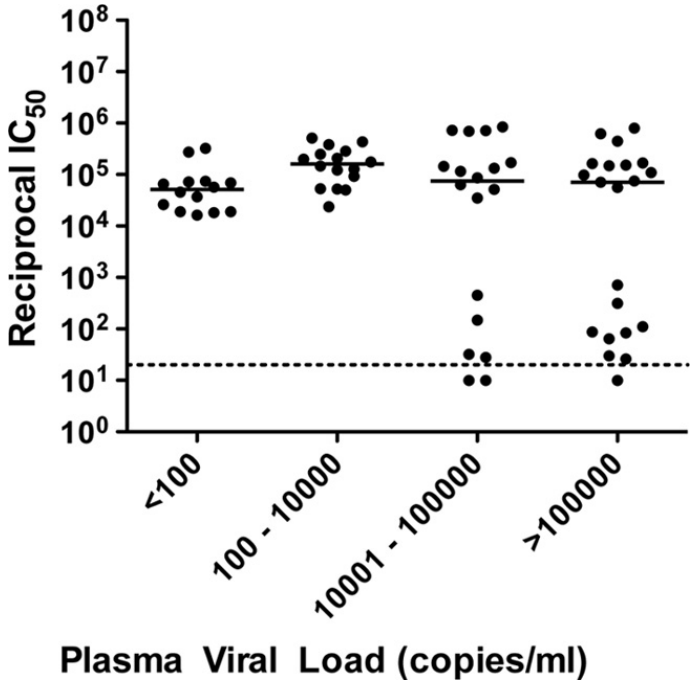


Figure 5.4 Contemporaneous autologous neutralisation of HIV-2 envelopes stratified by plasma viral load. Of the 69 primary HIV-2 envelopes (from 21 subjects) used, 64 represented gp140 cloned into 7312A-SNAG. Values are means of two or three independent experiments. No significant difference in median IC_{50} is seen between viral load categories ($p = 0.09$).

A subset of envelopes (n = 15/69) was, however, found to be relatively resistant (IC_{50} < 1:1000) to autologous contemporaneous neutralisation and these envelopes were isolated more frequently from patients with higher plasma viraemia ($p < 0.001$, two-tailed Fisher's exact test for the proportion of resistant clones if viraemia <10,000 *versus* ≥ 10,000 copies/ml). Three of these envelopes were included in the panel used in heterologous Nab studies and found to be either resistant to neutralisation by the majority of plasma samples tested (CA4205.8 and

CA4182.4, Figure 5.2), or display only limited neutralisation susceptibility (CA409.14, Figure 5.2). In addition, five (of seven) subjects from whom more resistant envelopes were isolated (TD15, TD16, TD19, TD20, TD21), also had high titre contemporaneous autologous Nab against other variants in the quasi-species. Taken together, this indicates that the lower Nab titres seen in some cases reflect envelopes with a neutralisation-resistant phenotype, rather than subjects with poorer host humoral responses. Autologous Nab titres against six envelopes (of seven used in both autologous and heterologous Nab studies) were 1.5 – 8-fold higher than the median heterologous titre against the same envelope, which would suggest that regardless of the neutralisation phenotype of the envelope, HIV-2 infected subjects retain the ability to mount effective Nab responses specific to Env epitopes displayed on contemporaneous autologous isolates. No significant association was seen between autologous and heterologous Nab titres in the same individual, although data from only 7 subjects were available for this analysis.

Neutralisation by anti-HIV-2 Env human monoclonal antibodies reveals diverse epitopes across the HIV-2 envelope are targeted in natural infection but are nevertheless susceptible to neutralisation escape

To characterise the neutralisation phenotype of the isolated HIV-2 Env variants further, a subset was tested against five anti-HIV-2 Env human monoclonal antibodies (Mabs) with varied specificities. Although there was heterogeneity in susceptibility to Mabs amongst the 21 primary envelopes (from 13 subjects) and the two reference strains tested (Table 5.2), those envelopes that were relatively resistant to autologous or heterologous plasma neutralisation, were globally resistant to neutralisation by all HIV-2 human Mabs at the highest concentration tested (10μg/ml). The anti-V3 Mab, 6.10F, showed the most potent neutralisation

and Mab 1.4H, thought to be a CD4-induced antibody and therefore potentially target the coreceptor binding site, neutralised many HIV-2 envelopes in the absence of soluble CD4. These data suggest that in HIV-2 natural infection, Nabs with a breadth of epitope specificities are elicited that can potently neutralise a variety of diverse envelopes, but also that neutralisation resistant HIV-2 envelopes are likely to represent variants that have accumulated mutations in multiple sites, resulting in a globally resistant phenotype.

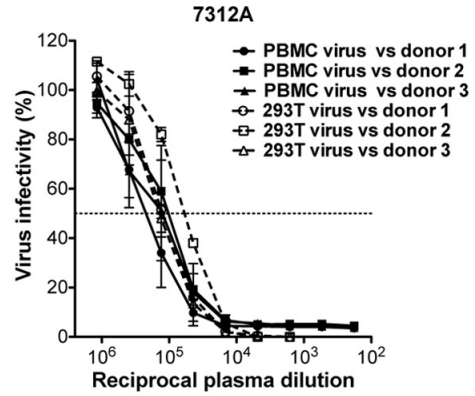
Table 5.2 Neutralisation of HIV-2 envelopes with 5 human HIV-2 monoclonal antibodies. Values given are mean IC_{50} (μg/ml) from two or three separate experiments. Values for neutralising antibody titres are reciprocal plasma autologous titres unless otherwise stated.

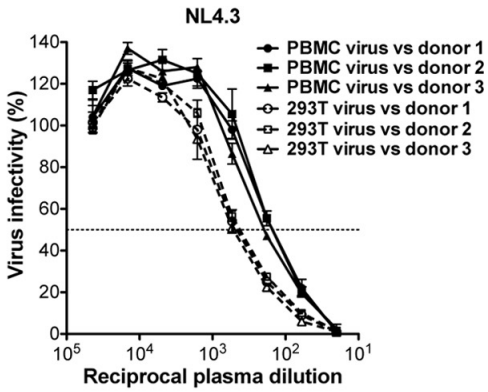
	1.7A (V4 dependent)	**6.10F (V3)**	**6.10B (CD4 BS)[1]**	**1.4B (overlaps V3)**	**1.4H (CD4i)[2]**	**Plasma neutralisation titre**
7312A	0.08	0.004	0.1	0.038	**>10[b]**	2.8×10^{5} [H]
ROD A [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**32[H]**
CA316.9	0.18	0.01	0.19	**>10**	**>10**	7.3×10^{5}
CA319.7	**>10**	0.015	0.26	0.05	0.02	5.1×10^{5}
CA330.5	**>10**	0.01	0.165	0.18	0.175	5.7×10^{4}
CA381.4	0.025	0.01	0.07	**>10**	0.4	4.4×10^{5}
CA4401.1	4.59	0.01	0.17	0.415	2.5	2.0×10^{5}
CA4153.9	**>10**	0.01	0.095	1.02	0.015	7.2×10^{5}
CA7205.8	0.015	0.0075	0.18	0.02	0.015	2.5×10^{5} [H]
CA4145.13	0.17	0.006	0.2	0.18	0.02	3.8×10^{5}
CA4291.2	0.54	0.006	0.155	0.115	0.025	1.6×10^{4}
CA409.14 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**64**
CA4205.1 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**713**
CA4205.8 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**87**
CA4137.6	**>10**	0.015	0.09	0.66	0.18	8.6×10^{4}
CA4137.2	**>10**	0.007	0.095	0.45	0.045	1.4×10^{5}
CA4137.19 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**<20**
CA4137.21	**>10**	0.01	0.07	0.3	0.035	1.7×10^{5}
CA4206.6	**>10**	0.03	0.65	0.44	0.1	5.6×10^{4}
CA4206.7	**>10**	0.01	0.2	0.6	0.17	1.7×10^{5}
CA4206.13	**>10**	0.04	1.01	0.95	0.21	9.8×10^{4}
CA4206.5 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**111**
CA4206.14 [R]	**>10**	**>10**	**>10**	**>10**	**>10**	**314**
VSV [V]	**>10**	**>10**	**>10**	**>10**	**>10**	-

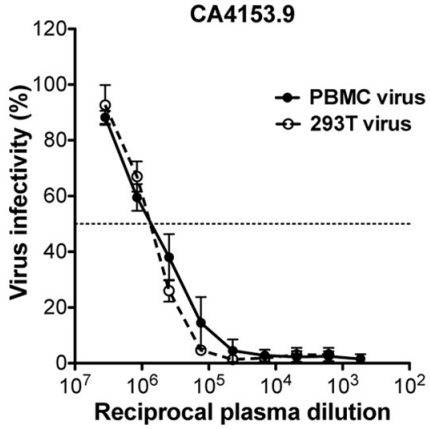
[1] CD BS, anti-CD4 binding site.
[2] CD4-induced.
[V] Vesicular Stomatitis Virus envelope.
[R] Envelope resistant to autologous and/or heterologous plasma neutralisation.
[H] Median reciprocal heterologous neutralising antibody titre (against 40 subject plasma samples).
[b] Bold data indicate envelope-Mab combinations where no neutralisation was observed.

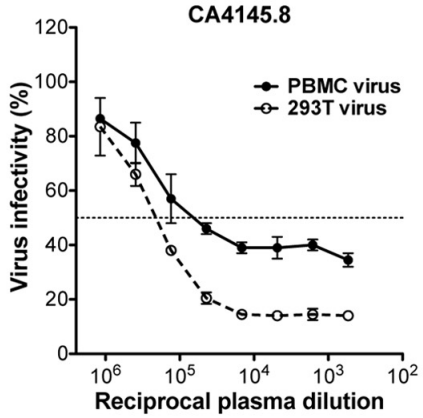
Virus passage in peripheral blood mononuclear cells variably affects neutralisation titres but not the overall HIV-2 envelope neutralisation-phenotype

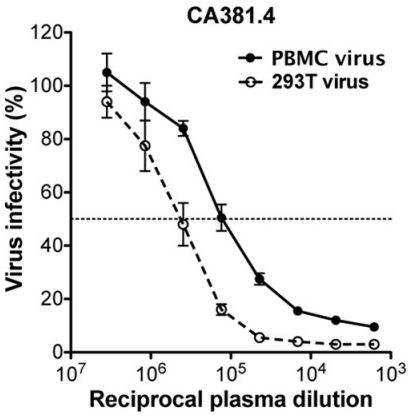
Differences in neutralisation sensitivity depending on the cell-type used for viral propagation have been previously noted in HIV-1 isolates [301]. The neutralisation phenotype of a subset of HIV-2 Env isolates (gp140 in 7312A-SNAG), grown in 293T cells described above, was evaluated following minimal passage in mitogen-stimulated PBMCs for 7 days. Along with 7312A and HIV-1 NL4.3, six primary HIV-2 Env isolates (of 14 tested) produced PBMC-passaged viral titres high enough after 7 days to assess autologous neutralisation in TZM-bl cells. Heterologous neutralisation of 7312A and NL4.3 was assessed using plasma from three randomly chosen HIV-2 and HIV-1 donors respectively. Consistent with previous observations [301], PBMC-passaged NL4.3 was less sensitive to neutralisation by all three donor plasma samples (Figure 5.5). However, no significant reduction in Nab titre of the highly sensitive HIV-2 Env isolate 7312A was observed. Likewise, no significant reduction in Nab titre of CA4153.9 was seen when PBMC-virus was assessed against contemporaneous autologous plasma. PBMC-CA381.4 and PBMC-CA4145.8 showed only modest reductions in neutralisation sensitivity. This indicates that the high Nab titres observed in HIV-2 infected subjects are unlikely to be due to an artifact of virus production in 293T cells. The neutralisation resistant isolates CA4137.19, CA4013.5 and CA4182.4 all remained resistant to (autologous contemporaneous) neutralisation following PBMC passage, with a clear indication that the PBMC-passaged viruses were even more resistant than their 293T-grown counterparts.

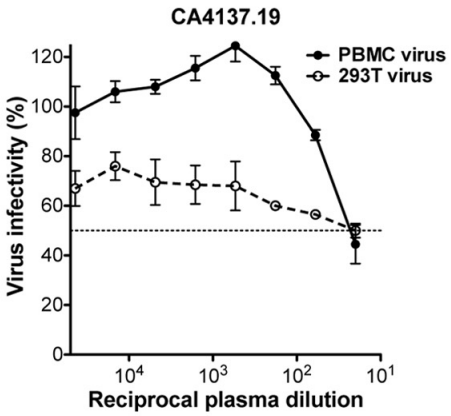

7312A
PBMC virus vs donor 1
PBMC virus vs donor 2
PBMC virus vs donor 3
293T virus vs donor 1
293T virus vs donor 2
293T virus vs donor 3
Virus infectivity (%)
Reciprocal plasma dilution

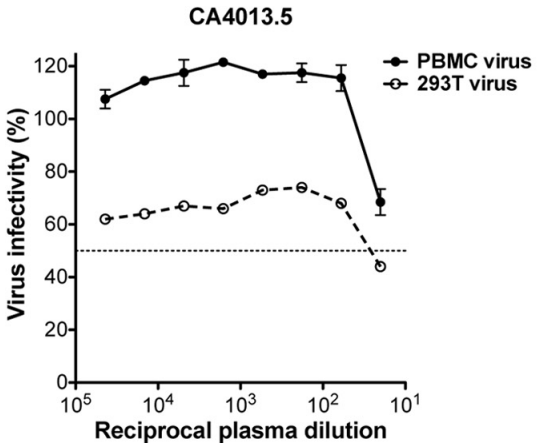

NL4.3
PBMC virus vs donor 1
PBMC virus vs donor 2
PBMC virus vs donor 3
293T virus vs donor 1
293T virus vs donor 2
293T virus vs donor 3
Virus infectivity (%)
Reciprocal plasma dilution


CA4153.9
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution


CA4145.8
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution


CA381.4
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution


CA4137.19
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution


CA4013.5
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution

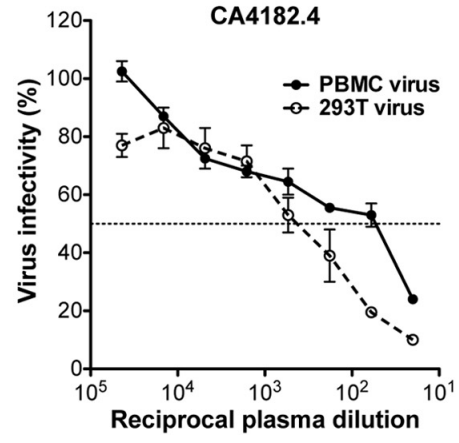

CA4182.4
PBMC virus
293T virus
Virus infectivity (%)
Reciprocal plasma dilution

Figure 5.5 Comparison of neutralisation of 293T-grown and PBMC-passaged virus with the same plasma sample. Shown are autologous neutralisation of CA4153.9, CA4145.8, CA381.4, CA4137.19, CA4013.5 and CA4182.4 and heterologous neutralisation of 7312A and NL4.3. Plotted values are means (and SDs) from 3 separate experiments. Virus infectivity (%) is compared to a virus-only condition.

Neutralisation resistance in HIV-2 is associated with greater env diversification and changes in N-linked glycosylation sites, but not coreceptor tropism or CD4-dependence.

To explore the relationship between intra-patient viral diversity and the presence of neutralisation-resistant envelopes, pairwise genetic distances between HIV-2 *env* isolates from each subject were calculated (Figure 5.6), using only those with ≥ 3 functional *env* variants available (n = 14). Subjects from whom resistant envelopes were isolated had significantly higher intra-patient diversity, when compared to those where only neutralisation-sensitive variants were found (median 0.034 *versus* 0.003 substitutions per site, p = 0.0007, two-tailed Mann-Whitney U test,). Of note, within each subject, no *env* variants were 100% identical at the nucleotide level. These findings also provide indirect support for an association between higher viraemia and viral diversification in HIV-2 infection, as the median plasma VL (range) in the group with higher diversity was 166,576 (43,396 – 320,272) compared to that in the group with lower intra-patient diversity 359 (<100 – 13,371) copies/ml (p = 0.002, two-tailed Mann-Whitney U test).

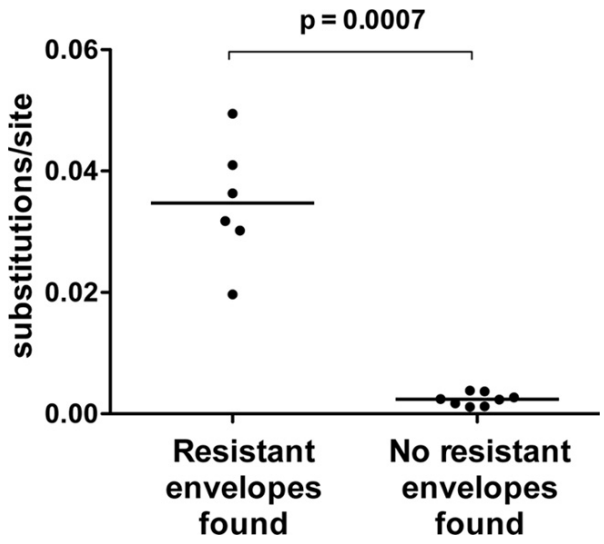
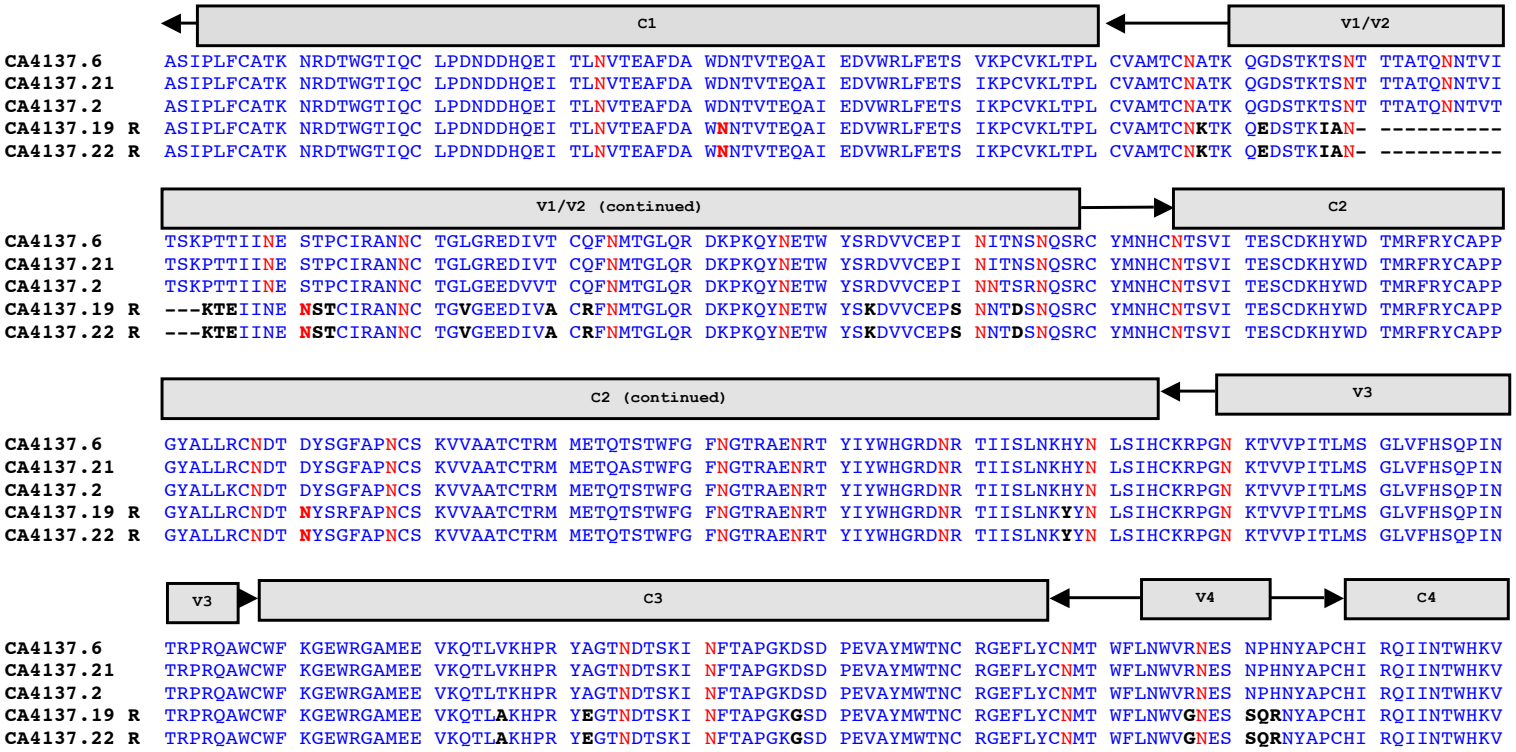


Figure 5.6 Association between greater intra-patient diversity and the presence of neutralisation-resistant HIV-2 envelopes. Plotted values are mean pairwise distances (substitutions/site) between *env* variants from each patient.

The concept of an evolving glycan shield as a mechanism of neutralisation escape is well described in HIV-1 infection [128]. Comparing sensitive and resistant Caió HIV-2 Envs used in this study, the number of potential N-linked glycosylation (PNLG) sites in any individual conserved or variable region or across the entire gp140 fragment (n = 26 median PNLG in both sets) was not significantly different between the two groups. It is likely however, that akin to HIV-1, escape from Nab may occur via multiple pathways and could be via a unique route within each HIV-2 infected subject [302]. PNLG sites were therefore compared in Envs from individuals where both resistant and sensitive variants had been isolated (n =5). In all individuals, differences in either the number or position of PNLG were observed between sensitive and resistant envelopes (e.g. TD15 and TD19, Figure 5.7a and b), although the extent of these differences varied between 1 – 4 additional or shifted sites. However, multiple additional amino acid substitutions unique to resistant

variants can be seen throughout the envelope ectodomain (including deletions in V1/V2 common to resistant variants in 2/5 individuals), thus highlighting the complexity of neutralisation escape in chronic HIV-2 infection.

5.7a


C1
V1/V2
CA4137.6 ASIPLFCATK NRDTWGTIQC LPDNDDHQEI TLNVTEAFDA WDNTVTEQAI EDVWRLFETS VKPCVKLTPL CVAMTCNATK QGDSTKTSNT TTATQNNTVI
CA4137.21 ASIPLFCATK NRDTWGTIQC LPDNDDHQEI TLNVTEAFDA WDNTVTEQAI EDVWRLFETS IKPCVKLTPL CVAMTCNATK QGDSTKTSNT TTATQNNTVI
CA4137.2 ASIPLFCATK NRDTWGTIQC LPDNDDHQEI TLNVTEAFDA WDNTVTEQAI EDVWRLFETS IKPCVKLTPL CVAMTCNATK QGDSTKTSNT TTATQNNTVT
CA4137.19 R ASIPLFCATK NRDTWGTIQC LPDNDDHQEI TLNVTEAFDA WNNTVTEQAI EDVWRLFETS IKPCVKLTPL CVAMTCNKTK QEDSTKIAN- ----------
CA4137.22 R ASIPLFCATK NRDTWGTIQC LPDNDDHQEI TLNVTEAFDA WNNTVTEQAI EDVWRLFETS IKPCVKLTPL CVAMTCNKTK QEDSTKIAN- ----------
V1/V2 (continued)
C2
CA4137.6 TSKPTTIINE STPCIRANNC TGLGREDIVT CQFNMTGLQR DKPKQYNETW YSRDVVCEPI NITNSNQSRC YMNHCNTSVI TESCDKHYWD TMRFRYCAPP
CA4137.21 TSKPTTIINE STPCIRANNC TGLGREDIVT CQFNMTGLQR DKPKQYNETW YSRDVVCEPI NITNSNQSRC YMNHCNTSVI TESCDKHYWD TMRFRYCAPP
CA4137.2 TSKPTTIINE STPCIRANNC TGLGEEDVVT CQFNMTGLQR DKPKQYNETW YSRDVVCEPI NNTSRNQSRC YMNHCNTSVI TESCDKHYWD TMRFRYCAPP
CA4137.19 R ---KTEIINE NSTCIRANNC TGVGEEDIVA CRFNMTGLQR DKPKQYNETW YSKDVVCEPS NNTDSNQSRC YMNHCNTSVI TESCDKHYWD TMRFRYCAPP
CA4137.22 R ---KTEIINE NSTCIRANNC TGVGEEDIVA CRFNMTGLQR DKPKQYNETW YSKDVVCEPS NNTDSNQSRC YMNHCNTSVI TESCDKHYWD TMRFRYCAPP
C2 (continued)
V3
CA4137.6 GYALLRCNDT DYSGFAPNCS KVVAATCTRM METQTSTWFG FNGTRAENRT YIYWHGRDNR TIISLNKHYN LSIHCKRPGN KTVVPITLMS GLVFHSQPIN
CA4137.21 GYALLRCNDT DYSGFAPNCS KVVAATCTRM METQASTWFG FNGTRAENRT YIYWHGRDNR TIISLNKHYN LSIHCKRPGN KTVVPITLMS GLVFHSQPIN
CA4137.2 GYALLKCNDT DYSGFAPNCS KVVAATCTRM METQTSTWFG FNGTRAENRT YIYWHGRDNR TIISLNKHYN LSIHCKRPGN KTVVPITLMS GLVFHSQPIN
CA4137.19 R GYALLRCNDT NYSRFAPNCS KVVAATCTRM METQTSTWFG FNGTRAENRT YIYWHGRDNR TIISLNKYYN LSIHCKRPGN KTVVPITLMS GLVFHSQPIN
CA4137.22 R GYALLRCNDT NYSGFAPNCS KVVAATCTRM METQTSTWFG FNGTRAENRT YIYWHGRDNR TIISLNKYYN LSIHCKRPGN KTVVPITLMS GLVFHSQPIN
V3
C3
V4
C4
CA4137.6 TRPRQAWCWF KGEWRGAMEE VKQTLVKHPR YAGTNDTSKI NFTAPGKDSD PEVAYMWTNC RGEFLYCNMT WFLNWVRNES NPHNYAPCHI RQIINTWHKV
CA4137.21 TRPRQAWCWF KGEWRGAMEE VKQTLVKHPR YAGTNDTSKI NFTAPGKDSD PEVAYMWTNC RGEFLYCNMT WFLNWVRNES NPHNYAPCHI RQIINTWHKV
CA4137.2 TRPRQAWCWF KGEWRGAMEE VKQTLTKHPR YAGTNDTSKI NFTAPGKDSD PEVAYMWTNC RGEFLYCNMT WFLNWVRNES NPHNYAPCHI RQIINTWHKV
CA4137.19 R TRPRQAWCWF KGEWRGAMEE VKQTLAKHPR YEGTNDTSKI NFTAPGKGSD PEVAYMWTNC RGEFLYCNMT WFLNWVGNES SQRNYAPCHI RQIINTWHKV
CA4137.22 R TRPRQAWCWF KGEWRGAMEE VKQTLAKHPR YEGTNDTSKI NFTAPGKGSD PEVAYMWTNC RGEFLYCNMT WFLNWVGNES SQRNYAPCHI RQIINTWHKV

5.7b

Figure 5.7 Alignments from two subjects (a and b) demonstrating amino acid changes specific to resistant (R) envelopes. Either additional or shifted potential N-linked glycosylation sites are highlighted in bold red, and other changes in black. Only regions with differences between resistant and sensitive envelopes are displayed for each subject.

Previous studies have also suggested an association between neutralisation sensitivity and CD4-independence [143], as well as resistance to neutralisation and CXCR4-tropism [303] in HIV-2. To explore whether the differences in neutralisation phenotype seen in the current set of primary HIV-2 envelopes, could be explained by similar mechanisms, their ability to infect NP2 cells expressing different combinations of CD4, CCR5 and CXCR4 was assessed (Tables 5.3). Thirty one HIV-2 envelopes (from 18 individuals) were tested, along with the X4-tropic HIV-1 isolate NL4.3, the CD4-dependent HIV-2 isolate ROD A and ROD B, a neutralisation-sensitive and CD4-independent variant produced via passage of ROD A in C8166 cells [304]. NL4.3, ROD A and ROD B all displayed their reported coreceptor usage. Only one primary envelope (CA65331.3) showed a dual R5/X4-tropic phenotype, whereas all other envelopes were R5-tropic. Consistent with previous data, ROD B was able to infect CD4-negative cells efficiently using CXCR4, but a similar ability to infect cells only expressing the primary coreceptor was not observed in any of the primary HIV-2 envelopes cloned directly from plasma viral RNA and propagated in 293T cells. An association between sensitivity or resistance to neutralisation and coreceptor tropism/CD4-dependence was therefore not evident in the HIV-2 envelope isolates described in this study.

Persistence of neutralisation sensitive envelope variants despite the presence of high magnitude autologous neutralising antibodies

In HIV-1 infection, constant viral escape from the autologous Nab response usually results not only in poor contemporaneous autologous Nab titres, but even lower Nab titres when Envs are tested against plasma samples from older time points [128], reflecting the inability of the Nab repertoire to keep up with *env* evolution under Nab pressure. One previous publication (using four subjects) has suggested

that, in contrast, escape from autologous Nabs is rarely seen in HIV-2 infection [142]. To extend this observation with a larger data set, available archived plasma samples from previous Caió cohort bleeds, were obtained for 10 subjects and the ability of these historical samples to neutralise one of the most recent neutralisation sensitive envelopes available (from either 2003 or 2006) was tested. In almost all cases, the autologous Nab titres in plasma samples taken up to 15 years previously, against HIV-2 envelopes from the most recent time point, were as high as those seen in contemporaneous samples (Figure 5.8a). The one exception was an approximate $1\log_{10}$ reduction in Nab titre of 1991 plasma against CA316.9, and in this case a similar reduction was seen in heterologous neutralisation of 7312A.

Tables 3a and 3b. Envelope coreceptor use and relationship to neutralisation phenotype.

3a

Subject	Envelope ^	Coreceptor use					Classification\	Neutralisation phenotype\\
		Infectivity on NP2 cells (10^3 FFU/ml)*						
		CD4/CCR5	CD4/CXCR4	CD4 alone	CCR5 alone	CXCR4 alone		
TD02	CA4258.11	81.5	-	nt	0.1	nt	R5	S
TD05	CA370.16	20.4	-	nt	-	nt	R5	S
TD06	CA4401.1	53	-	nt	-	nt	R5	S
TD08	CA381.4$	255	0.1	1.2	-	nt	R5	S
TD09	CA4145.8	627	-	nt	0.8	nt	R5	S
TD11	CA4272.16	330	0.2	1.5	-	nt	R5	S
	CA65331.3	110	50.0	0.8	-	nt	R5/X4	S
TD12	CA65332.1	42.5	-	nt	0.1	nt	R5	S
TD13	CA4153.9	2,500	2.0	1.3	0.9	nt	R5	S
TD15	CA4137.2	1,328	-	nt	-	nt	R5	S
	CA4137.19	7,125	1.4	1.7	1.2	nt	R5	R
TD16	CA4182.4	273	-	nt	-	nt	R5	R
	CA4182.12	4,687	1.0	0.9	1.1	nt	R5	S
	CA65410.3	225	0.2	0.8	-	nt	R5	S
	CA65410.5	134.5	0.1	1.5	0.5	nt	R5	S
	CA65410.7	300	0.4	2.5	0.9	nt	R5	S
	CA65410.13	59.5	-	nt	-	nt	R5	S
TD17	CA4205.1	850	2.5	2.1	1.1	nt	R5	R
	CA4205.8	1,100	5.5	3.0	1.4	nt	R5	R
TD19	CA4269.2	1,781	7.5	6.3	0.2	nt	R5	S

	CA4269.6	2,625	-	nt	1.1	nt	R5	R
	CA65339.3	32	-	nt	-	nt	R5	R
TD20	CA4206.7	25	-	nt	nt	nt	R5	S
TD21	CA4013.5	3,187	2.4	1.9	0.9	nt	R5	R
	CA4013.8	22.5	-	nt	nt	nt	R5	S
NA	NL4.3	-	212.5	nt	nt	nt	X4	NA

3b

	Envelope^^	Infectivity on NP2 cells (TCID50/ml)**						
		CD4/CCR5	**CD4/CXCR4**	**CD4 alone**	**CCR5 alone**	**CXCR4 alone**		
NA	ROD A	69,877	31,250	250	112	112	R5/X4	R
NA	ROD B	1,250	31,250	50	<50	13,975	X4/CD4-indep	S
TD01	CA330.5	69,877	559	559	250	112	R5	S
TD08	CA381.4[$]	69,877	250	112	250	112	R5	S
TD14	CA316.9	349,386	1250	559	559	559	R5	S
TD17	CA319.7	349,386	2795	559	559	559	R5	S
TD18	CA409.14	156,250	1250	1,250	1,250	1,250	R5	R
NA[§]	CA7205.8	349,386	2795	112	1,869	250	R5	S

nt Not tested.

^ HIV-2 gp140 from subjects cloned into full-length HIV-2 7312A-SNAG vector. NL4.3 refers to virus derived from the full-length HIV-1 NL4.3 molecular clone.

^^ HIV-2 gp160 cloned in to expression vector pcDNA3.1V5-His-TOPO. Pseudovirus produced by co-transfection with NL4.3LucR$^-$E$^-$.

* FFU – focus forming units measured by immunostaining of NP2 cells +/- CD4 +/- coreceptors following infection with live virus.

** TCID50 – Tissue culture infectious dose 50 calculated according to luciferase readout in NP2 cells following infection with pseudovirus produced by co-transfection of HIV-2 gp160 envelope plasmid + NL4.3LucR$^-$E$^-$.

\ Viruses were classified as dual-tropic if the titre on alternative coreceptor was within 1.5log$_{10}$ of the titer obtained from the dominant coreceptor cell line [305].

\\ S – sensitive R – resistant

$ Coreceptor tropism tested in both formats.

§ Subject ID not assigned as envelope not used for autologous neutralisation studies.

Although the main objective of this study was not a longitudinal assessment of contemporaneous Nab responses, three individuals had adequate plasma available from two time points to isolate envelopes from both 2003 and 2006 (TD11, TD16 and TD19; Table 4.1). Two subjects (TD16 and TD19) had both sensitive and resistant envelopes isolated from 2003 plasma (Figures 5.8c and 5.8d). For subjects TD11 and TD16 (Figure 5.8b and 5.8c respectively), no escape from neutralisation was observed between envelopes isolated in 2003 and 2006. In fact, despite the presence of a relatively resistant envelope isolated in 2003 for subject TD16, surprisingly all four 2006 envelopes were sensitive to autologous neutralisation (Figure 5.8c). In contrast, subject TD19 had one sensitive and two resistant Env variants isolated from 2003 plasma, but all 2006 envelopes were neutralisation-resistant, indicating possible escape (Figure 5.8d).

Selective pressure on the HIV-2 envelope in the face of potent autologous neutralising antibody responses

The neutralising antibody response is well recognised in HIV-1 infection as the major driving force in *env* evolution and positive selective pressure [306], with resulting continual escape from Nab. Previous studies, focusing on the C2V3C3 region of the HIV-2 envelope, have suggested *env* evolution is slower in asymptomatic infection [101], as well as highlighting that this region in HIV-2 is largely under purifying or negative selective pressure, implying that possible functional constraints may play a role [273, 307]. In view of the exceptionally high autologous Nab observed in the current study, as well as the low neutralisation escape seen, selective pressure was examined across the entire HIV-2 gp140 in the set of 70 functional envelopes that were isolated in this study. Using the three different algorithms [297, 298], the number of sites under significant positive

selective pressure was remarkably low under SLAC (11 sites) and FEL (19 sites), although higher under the FEL algorithm (44 sites); with the number of negatively selected sites being 175, 232 and 346 respectively (from a total of 641 codons). To compare how this differs from diversification of, and selective pressure on, the HIV-1 envelope in the same community, available partial C2 to gp41 ectodomain sequences from 75 HIV-2 and 56 HIV-1 infected individuals from Caió were used; a fragment corresponding to nucleotides (nt) 6584 – 7490 of HIV-1 HXB2 (K03455) and nt 6984 – 7859 of HIV-2 ROD (M15390). Phylogenetic studies and the REGA HIV subtyping algorithm [232] showed all HIV-2 sequences belonged to group A and all HIV-1 sequences used were subtype CRF02_AG (See Chapters 2 and 3). Surprisingly pairwise genetic distances in the HIV-2 set were significantly higher than in HIV-1, revealing a greater diversity in HIV-2 at the nucleotide level in this community (Table 5.4). This is likely to reflect the different ages of the two epidemics in Guinea-Bissau, with HIV-1 having been introduced more recently than HIV-2 [51]. Despite this greater diversity in HIV-2, amino acid entropy and the number of sites under positive selection was significantly higher in the HIV-1 envelope whereas the number of negatively selected sites was much higher in the HIV-2 envelope (Table 5.4 and Figure 5.9). It appears, therefore, that at least at a community level there are constraints present on the HIV-2 envelope that restrict amino acid changes when compared to HIV-1, despite substantial pressure from neutralising antibody responses.

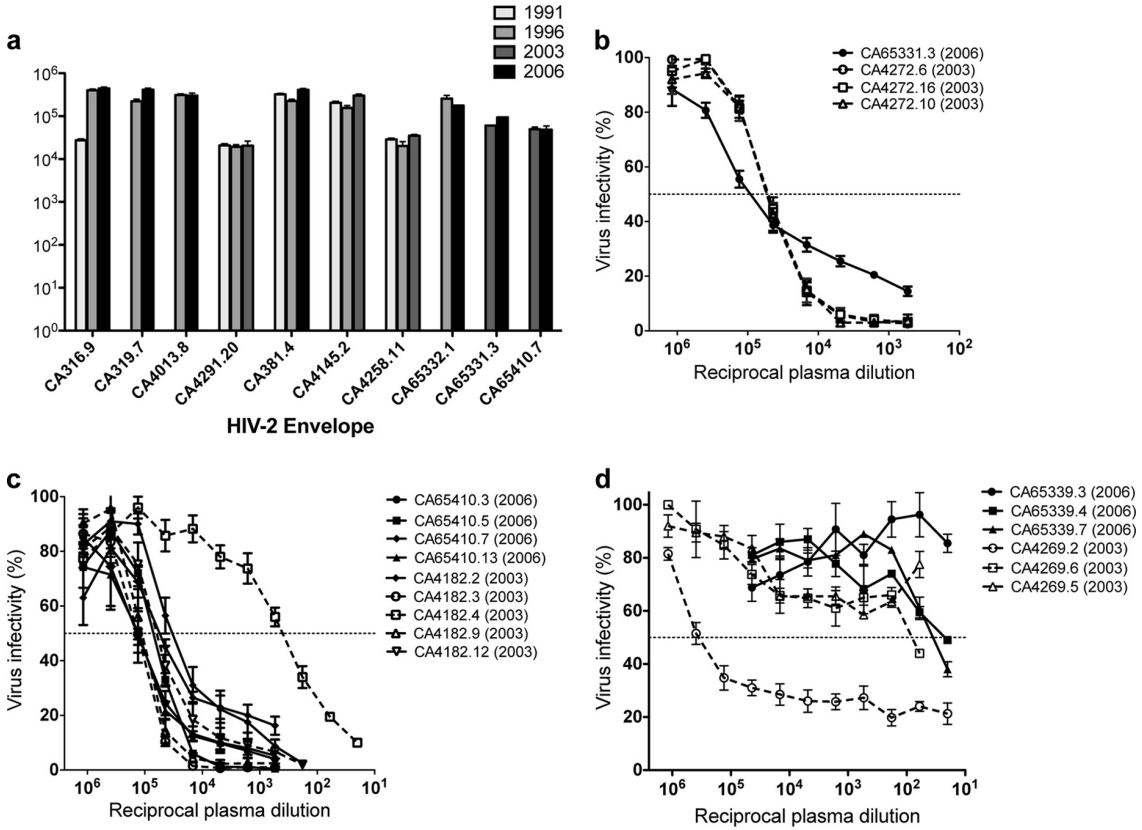


Figure 5.8a – d. Assessment of neutralisation escape in HIV-2. (a) Neutralisation of HIV-2 envelopes by autologous plasma from older time points in 10 subjects. The legend depicts the year of plasma sampling. HIV-2 envelopes were isolated from the most recent time point of plasma collection depicted for each clone (either 2003 or 2006). (b to d) Neutralisation of envelopes from either 2003 or 2006 by

contemporaneous plasma for subjects TD11 (b), TD16 (c), and TD19 (d), from whom variants were isolated at both time points. All values represent means (and SDs) of three independent experiments. Virus infectivity (%) is compared to a virus-only condition.

Table 5.4 Comparison of HIV-2 and HIV-1 envelope (partial C2 – gp41 ectodomain) diversity, amino acid entropy, and sites under positive and negative selective pressure.

	HIV-2	HIV-1	p-value
Mean pairwise genetic distance - substitutions/site	0.183 (0.181 - 0.185)	0.148 (0.146 - 0.151)[a]	$p < 0.0001$
Mean Shannon's entropy	0.255 (0.207 - 0.303)	0.347 (0.292 - 0.402)	$p = 0.01$
Selection analysis (from a total 277 codons)[b]:			
SLAC			
Positively selected sites	5	19	$p < 0.0001$
Negatively selected sites	122	53	
FEL			
Positively selected sites	11	22	$p < 0.0001$
Negatively selected sites	140	79	
REL			
Positively selected sites	21	51	$p < 0.0001$
Negatively selected sites	195	143	

a Data in parentheses are 95% confidence intervals. *b* Codons with a P value of <0.05 (SLAC and FEL) or a Bayes factor of >50 (REL) considered significant.

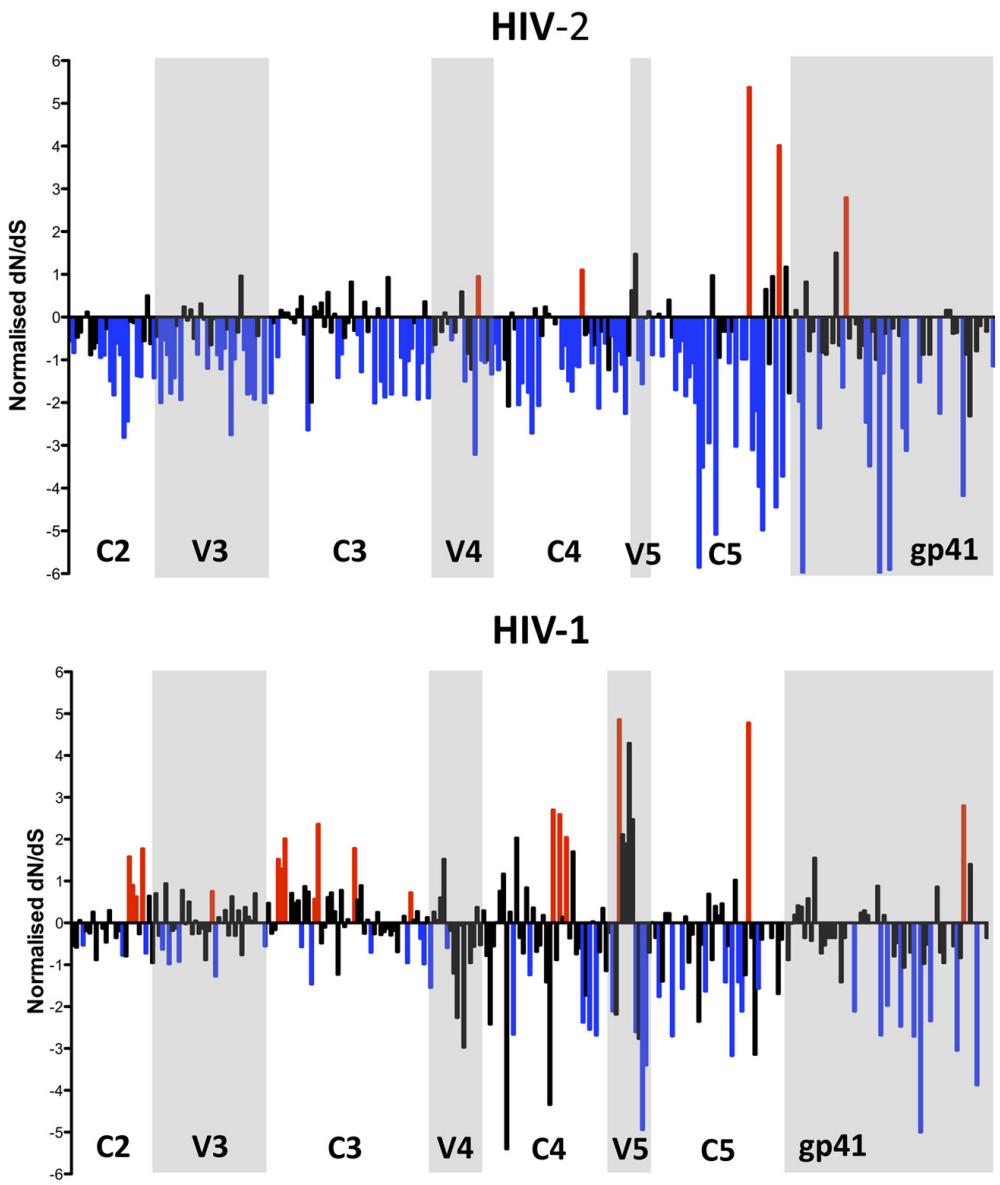


Figure 5.9 Comparison of the ratio of nonsynonymous (*dN*) to synonymous (*dS)* changes in HIV-2 (n = 75) and HIV-1 (n = 56) partial C2 to gp41 ectodomain fragments. Plotted values are those obtained under the SLAC algorithm with values normalised to alignment length. Positions found to be under significant positive selection (red) and negative selection (blue) under all three algorithms (SLAC, FEL, REL) are highlighted.

5.4 DISCUSSION

A major obstacle to the development of an antibody-based HIV-1 vaccine over the last 25 years has been identifying what defines protective humoral immunity in HIV infection: that is, lack of clarity over the role, if any, of Nabs elicited during natural HIV infection in limiting viral replication and the course of disease progression. Previous studies have reached contrasting conclusions. The constant antibody-driven viral evolution in HIV-1 *env* [129] indicates that the Nab response must be sufficiently functional to induce viral change. While some studies suggest a role in viral control [308, 309], most studies (both in acute and chronic HIV-1 infection) have failed to reveal a clear relationship between potent Nabs and lower VLs or slower disease progression [135-138]. With a high proportion of HIV-2 infected individuals with low or undetectable VLs, HIV-2 infection offers an important example of viral control with which to explore the relative importance of different host and viral factors implicated in limiting HIV pathogenesis. Only a handful of studies have attempted to establish the relationship between Nab in HIV-2 infected individuals and disease outcomes [141, 142, 145], in part due to the technical limitations in culturing primary virus or generating plasma RNA-derived *env* clones from individuals with low VLs, with which to measure autologous Nabs. Along with a study by Kong *et al.* [310] published concurrently with the current work, this study is one of the first to use cloned *env* plasma variants in a luciferase reporter gene assay to describe both autologous and heterologous Nab responses in HIV-2 infection; and with the inclusion of subjects ranging from LTNPs with undetectable plasma VL to HIV-2 progressors with high VL, it is also arguably the most comprehensive to date.

Heterologous Nab titres of strikingly high magnitude (IC_{50} 1:7000 – 1000000), several log_{10} higher than what is usually observed in HIV-1 infection using the same Nab assay [136], were found against primary HIV-2 envelopes in all 40 subject plasma tested. Yet despite the presence of potent heterologous Nabs, a significant positive association was found between these titres and plasma VL. This is consistent with findings from the one previous study on HIV-2 using a pseudotype reporter assay including only asymptomatic subjects [145] and confirms that similar to HIV-1 infection [308], greater antigenic stimulation and/or abnormal polyclonal B-cell activation in HIV-2 progressors may result in greater heterologous Nab production. If Nab do play a role in limiting HIV-2 replication, however, it is more likely to be reflected in contemporaneous autologous Nabs, which are conspicuously low or absent in most subjects with chronic HIV-1 infection [308, 311]. Exhibiting a clearly different scenario from that in HIV-1-infected individuals, equally potent (IC_{50} >1:10000) contemporaneous autologous Nabs were found in 78% of plasma-envelope combinations tested, including in those with undetectable plasma VLs. HIV-1 elite controllers, in contrast, have very low autologous Nab titres [135, 138], leading to the argument that antigenic stimulation is required to maintain Nab. It is possible that this unique feature in HIV-2 elite controllers is due to factors such as intermittent viral blips (which are difficult to detect with the Caió cohort design [51, 75]) providing a stimulus or ongoing viral replication in lymphoid tissue sufficient to maintain Env-specific memory B-cells. It is worth noting that despite plasma viraemia being below the limit of detection (100 copies/ml) in these individuals, the ability to amplify *env* using nested PCR in itself supports the existence of low level circulating viraemia; potentially priming and in turn being controlled by adaptive immune responses. Previous work in the same cohort has also observed a proportion of individuals

with viraemia below 100 copies/ml have detectable RNA by qualitative PCR [76]. Interestingly, despite the absence of detectable plasma viraemia, HIV-2 elite controllers are also able to maintain a remarkably high proportion of HIV-2 specific CD8+ T-cells in their circulation many decades after initial infection [172, 174] (and Chapter 6). However, despite the presence of potent contemporaneous autologous Nab, a significant inverse relationship between Nab titres and plasma viraemia could not be demonstrated. Given the 30-fold lower viraemia in HIV-2 infection compared with HIV-1 infected individuals in the asymptomatic stage [93], it is tempting to speculate that these high magnitude Nab could still play a role in containing HIV-2 replication, but what is clear from the current study is that they are certainly not the predominant force responsible for the dichotomous outcomes observed in HIV-2 infection. In fact, if only Nab titres against sensitive envelopes are considered, individuals with detectable (≥ 100 copies/ml) plasma VL do have significantly higher median autologous Nab titres than those with VL <100 copies/ml (1.47×10^5 vs. 5.1×10^4, $p = 0.0006$). This suggests that although autologous Nab titres are unusually high in most HIV-2 infected subjects, a relationship between greater antigenic load and higher magnitude autologous Nab may exist in HIV-2 infection; similar to the findings in heterologous Nab, as well as previously reported findings in HIV-1 infection [135].

An interesting finding was the presence of HIV-2 envelopes that displayed a more resistant neutralisation profile, notably with contemporaneous autologous titres (IC_{50} <1:1000) closer to what is observed in HIV-1 infection. Of note, most of the isolated HIV-2 envelopes were either hyper-susceptible or very resistant to neutralisation, akin to HIV-1 tier 1A and tier 3 envelopes respectively; whereas the majority of primary HIV-1 envelopes described tend to belong to the intermediate-

susceptible tier 2 category [312]. Although this finding requires confirmation with a greater number of individuals, the appearance of more resistant envelopes with higher viraemia might suggest a scenario where HIV-2 replication to a level where neutralisation escape is possible occurs more frequently with progressive HIV-2 infection. The association of greater intra-patient viral diversity with the presence of these envelopes would support this hypothesis, given that the capacity HIV-1 has for diversification underlies its ability to escape readily from host immunity. The persistence of neutralisation sensitive envelopes, despite the presence of Nab pressure for up to 15 years previously, also provides indirect support for the idea that neutralisation escape occurs less often in HIV-2 infection [142]. Certainly a previous report has highlighted that in asymptomatic HIV-2 infection, long-term *env* evolution is remarkably low [101], although the authors suggest that lack of Nab pressure may account for this observation. The current findings make an alternative explanation necessary. In some cases, these envelopes were found together with neutralisation sensitive variants in the same individual, an observation mirrored in some HIV-1 studies [135, 313-315]. As evidence from acute HIV-1 infection shows constant neutralisation escape usually results in successive replacement of neutralisation sensitive variants with waves of variants resistant to the autologous antibody response [128, 316], this is a curious finding. One explanation might be that most neutralisation sensitive variants are derived from a long-lived latent viral reservoir, thought to be a prominent feature of HIV-2 infection and inferred due to a lack of phylogenetic separation between HIV-2 sequences over time [101]. Nevertheless, in the face of greater HIV-2 replication, one would still expect the selective pressure exerted by such high Nab titres to result in elimination of neutralisation sensitive variants. In contrast, in one subject (TD16), a neutralisation resistant Env present (amidst several sensitive variants)

in 2003, failed to become the predominant population three years later. It is possible, therefore, that some sensitive HIV-2 Envs offer a selective advantage for alternative reasons. Extending previous studies focusing on the HIV-2 C2V3C3 [273, 307], there was limited evidence of positive selection across gp140 and significantly greater negative selective pressure when compared to HIV-1 in the region corresponding to C2 – gp41. Together, these allude to the possibility that certain functional constraints unique to the HIV-2 envelope may exist that limit amino acid changes, exerting a greater selective pressure than that of autologous Nab. Whether this is related purely to replicative fitness or additional HIV-2 Env functions such as tetherin antagonism [84] requires further investigation.

It is difficult to conclude whether the exceptionally high Nab titres seen in most envelope-plasma combinations are due to superior Nab production by HIV-2 infected individuals or features of most HIV-2 envelopes that result in greater susceptibility to Nab. Lack of *env* evolution and persistence of Nab epitopes over long periods of time may allow greater affinity maturation by B-cells that is not possible in HIV-1 infected individuals and CD4+ T-cell help is also better preserved in asymptomatic HIV-2 infection [169]. A more open conformation and exposed coreceptor site has been proposed for the HIV-2 envelope, as CD4-independent infection is related to global neutralisation sensitivity in some HIV-2 isolates [143]. There was, however, no evidence of this phenotype in the current set of primary envelopes; although interestingly the Mab 1.4H, thought to be a CD4-induced (CD4i) antibody (James Robinson, personal communication), was able to neutralise many isolates in the absence of soluble CD4 (sCD4). Previous work has shown that several HIV-2 strains (including 7312A) are only neutralised by HIV-1 CD4i Mabs in the presence of sCD4 [144], suggesting that an accessible coreceptor

site is not a typical characteristic of the HIV-2 Env. Elegant work from the same group has also shown that transplantation of HIV-1 V3 in the HIV-2 envelope not only allows detection of potent anti-V3-Nab in HIV-1 sera, but also results in a significant reduction in neutralisation by HIV-2 sera [317]; implying that anti-HIV-2 V3 responses contribute to Nab in HIV-2 sera and that this region may be particularly exposed in the native HIV-2 envelope. A recent publication using structural modeling has claimed, however, that the HIV-2 V3 loop is much *less* exposed than in HIV-1 [273]. Although the data showing a high degree of conservation in HIV-2 V3 are consistent with this hypothesis, the finding that the anti-V3 Mab 6.10F produced the most potent Mab neutralisation would favour the conclusions drawn by Davis *et al.*[317], and imply the lack of entropy in HIV-2 V3 may again be due to particular structural or functional constraints that limit its evolution. Although CXCR4-using HIV-2 isolates have been described in the literature [90, 303], a recent report suggests that R5 to X4 switch may occur less often than in HIV-1 [318]. The observation that only one of 31 primary envelopes tested (including many from HIV-2 progressors) was able to use CXCR4 efficiently for cell entry would support this finding and may also relate to an evolutionary constraint on the V3 region, which is well known to be the major determinant of coreceptor tropism.

Although there was no evidence that CD4-independence or CXCR4 usage was associated with neutralisation resistance in the set of primary envelopes, comparison of resistant and sensitive envelopes found within the same individual revealed clear changes in N-linked glycosylation sites within the resistant variants, which are known to mediate escape from neutralisation in HIV-1 infection [128, 314, 319]. These changes in PNLG, as well as other amino acid mutations common

to intra-patient resistant variants were found across the entire ectodomain, but were concentrated largely in V1/V2. This is in agreement with studies in both acute and chronic HIV-1 that suggest V1/V2 may be a major target of Nabs and changes within this region are observed during neutralisation escape [137, 319-322]. Despite the observation that all envelopes resistant to plasma neutralisation were globally resistant to all Mabs tested, no clear patterns were observed in PNLG changes or other mutations across these isolates from different individuals that would reliably predict the resistant phenotype. This is perhaps not surprising given the complexity in neutralisation escape described in HIV-1 with resistance to a single antibody possible through different pathways [302]. Thus this sequence-based comparison reveals parallels that exist between HIV-1 and HIV-2 when the latter infection does result in *env* evolution and escape from Nab, but further detailed experimentation and inclusion of more individuals is required to expand on these findings.

This study has certain limitations that must be mentioned. Firstly, all the HIV-2 envelopes were derived through bulk PCR and molecular cloning, which would potentially have been at risk of creating recombinants that do not exist *in vivo* through Taq-polymerase driven template switching [323]. Single genome amplification (SGA) techniques are now increasingly used to avoid such issues [191], but at the time were beyond the capability or resources of the current study. More recent work, however, found no difference in sampling bias between standard bulk PCR/cloning and SGA, provided an adequate number of templates are analysed [324]. Preliminary data (not shown) suggest that in subjects with low or undetectable plasma HIV-2 viraemia, the amplification conditions used are close to those that fulfil criteria for SGA (<30% PCR positivity[191]), but it may be

important to confirm these findings through formal SGA analyses; particularly to demonstrate the presence of the HIV-2 envelopes with varying neutralisation sensitivity. Of note, recent work using *envs* derived by SGA has indeed observed the presence of neutralisation sensitive and resistant HIV-2 envelopes from a single individual [310]. Secondly, for the heterologous Nab studies, individuals were selected who were not known to be epidemiologically linked to those from whom envelopes were derived (which could potentially result in Nab titres appearing high purely through shared epitopes due to shared transmission). Given the cohort setting, it is still inevitable that many infected individuals have viruses that share a recent common ancestor. However, as similarly high titres were observed in Caió plasma against 7312A, an isolate originally from Côte d'Ivoire, and in Gambian plasma samples against Caió HIV-2 envelopes (data not shown), it is unlikely that the high antibody titres seen are unique to this cohort alone. Furthermore, recent findings from Kong *et al.* [310] using HIV-2 infected sera predominantly from Senegal and Côte d'Ivoire and an identical neutralisation assay show equally potent Nab titres.

Despite the widespread use luciferase reporter gene assays with TZM-bl cells as targets in HIV-1 studies, clear differences exist in assay sensitivity and performance when compared to other available neutralisation assays [325]. Although is clear that the high neutralisation titres seen in HIV-2 sera is not an artifact of virus propagation in 293T cells, it would be important to validate these findings in the future using a range of different target cells and read-outs. Encouragingly, recent work has shown similar anti-HIV-2 Env titres using a PBMC neutralisation assay [310]. Little is known regarding the degree of cell-cell spread of HIV-2 and previous work has shown that spread of HIV-1 in this manner can be

inhibited by Nab [326]. Future work should attempt to establish whether anti-HIV-2 Nab can also prevent cellular HIV-2 transmission via the virological synapse and if the presence of these specific Nabs correlate with control of viraemia.

Although the analyses of *env* diversity and selection pressure at a population level provide intriguing insight into HIV-2 pathogenesis, more detailed studies comparing intra-patient sequences in individuals with varying disease outcomes could be considered. Of particular interest would be an analysis of whole genome sequences to determine whether these findings are generalisable across the entire HIV-2 genome. Notably, escape within CTL epitopes in HIV-2 infection are also yet to be described, despite the presence of potent Gag-specific CTL responses [174]; which is surprising given the frequency with which this occurs in HIV-1 [327, 328]. Finally, a study of Nab development and maintenance during and after acute HIV-2 infection would definitively resolve many unanswered questions about the role of Nab in defining the initial course of infection and whether escape from Nab is limited. However, the reducing incidence of HIV-2 infection on a background of substantially lower transmission than HIV-1 makes this incredibly difficult [51, 52].

In conclusion, the current study reveals remarkably high autologous and heterologous Nab titres in natural HIV-2 infection (and seemingly rare escape from Nab), yet these Nabs do not appear to play a major role in determining the level of HIV-2 viraemia. One could question what implication this has for predicting the benefit achievable through future Nab-based HIV vaccines, but of course the nature and titres of Nab required for prophylaxis and that necessary for viraemic control once infection is established may be very different. It is conceivable that the Nab titres seen in HIV-2 infection prevent further super-infection with other HIV-2

strains; this is not an uncommon occurrence in HIV-1 (reviewed in [329, 330]) but has not been extensively described or investigated in HIV-2. The role of other antibody-dependent effector functions such as complement-mediated lysis [331] or antibody-dependent cellular toxicity [332] have not been described in HIV-2 infection and warrant investigation within this model. Data from the current study also adds to mounting evidence that the potential for adaptive selection and escape from host immunity may be limited in HIV-2 infection, which may underlie the lower VLs and slower disease progression observed in HIV-2 infected individuals. Determining whether this is due to a yet unconfirmed potent host immune response or an intrinsic property of the virus deserves further investigation and has the potential to aid the understanding of HIV-1 pathogenesis and search for novel therapeutic targets considerably.

CHAPTER 6: T-CELL RESPONSES ASSOCIATED WITH HIV-2 VIRAL CONTROL: INSIGHTS INTO NATURAL CONTAINMENT OF A HUMAN RETROVIRAL INFECTION.

6.1 INTRODUCTION

Nearly three decades after HIV-1 was identified as the causative agent of AIDS, more than 33 million people worldwide are thought to be infected with the virus. Considerable progress has been made in understanding the mechanisms by which HIV interacts with human cells, evades immune recognition and profoundly damages the immune system, but a vaccine that reliably prevents or controls HIV infection remains elusive. Important insights into correlates of protective immunity may be gained from studies of SIV in animal models and human subjects infected with HIV-1 who do not develop immunodeficiency. One such model that may not have been adequately investigated is human infection with HIV-2. Although establishing the mechanisms of protection from disease progression in natural SIV hosts has implications for understanding how to control HIV-1 (reviewed in [333]), arguably, deciphering natural HIV-2 control in humans is a far more relevant comparison. Whereas prolonged survival with SIVsmm in the natural host is associated with high levels of viral replication, low levels of immune activation and weak immune responses, what is essentially the same virus behaves very differently in humans (reviewed in [334]). Non-progression with HIV-2 is associated with strong cellular immune responses [172] and undetectable viral load.

HIV-1 and HIV-2 share 30-60% sequence identity and show many similarities in their cytopathic potential and replication kinetics *in vitro* (reviewed in [334]). However, HIV-2 differs from HIV-1 in that a substantial proportion of infected people (estimated at 37% in the Caió community HIV-2 cohort) maintains an undetectable plasma viral load (VL, <100 copies/ml) for a decade or more and shows no signs of immunodeficiency [281]. Unlike the rare phenomenon of "elite control" in HIV-1 infection, this phenotype is both clinically stable and occurs more commonly than disease progression. Nevertheless, those people with HIV-2 who develop disease (15-20% in the Caió cohort) do so in a manner [62] and timeframe [71] that cannot be distinguished clinically from HIV-1 infection. Identifying the key differences between these two groups at either end of the clinical spectrum of HIV-2 infection should provide insights into the mechanisms by which humans can control infection with a potentially pathogenic retrovirus [335].

Previous studies have shown that HIV-2-infected people with a normal CD4+ count are more likely to preserve a polyfunctional virus-specific CD4+ T-cell response than their HIV-1-infected counterparts [169, 170], but whether these polyfunctional T-cell responses distinguish HIV-2 elite controllers (ECs) from viraemic progressors is not known. When the T-cell response towards the HIV-2 proteome was compared between ECs and viraemic patients, viral control was strongly associated with high magnitude responses against HIV-2 Gag peptides [172]. This study highlighted the role of T-cell responses to conserved epitopes in Gag p26, but was performed using ELISpot assays that did not distinguish between CD4+ and CD8+ responses as the major correlate of viral control. This chapter represents a detailed study of HIV-2 Gag-specific T-cell responses from subjects carefully selected to represent the clinical extremes of the Caió cohort, in order to

identify unequivocally the cellular immune correlates of HIV-2 viral control. This study arguably represents the most comprehensive analysis of HIV-2-specific T-cells to date, including detailed *ex vivo* phenotypic analysis of CD8+ T-cells targeting several immunodominant epitopes and direct comparison of both cytokine-mediated and cytotoxic functional capacity in HIV-2 ECs and viraemic subjects.

6.2 METHODS

Laboratory and flow cytometry analysis methods are detailed in Chapter 2, sections 2.16 – 2.22. Methods for amplification of the HIV-1 capsid are detailed in Chapter 2, section 2.3.

HIV-1 and HIV-2 capsid sequence analysis and tests for codon selection

Sequences were aligned using CLUSTALW2 [220, 221] and all alignments inspected and manually edited using Se-Al (Sequence Alignment Editor, v2.0a11) [222]. Ambiguous regions were deleted prior to further analyses. The GTR+I+G model of evolution was selected as the best-fit model of substitution for both HIV-1 and HIV-2 capsid alignments by jModelTest [293]. HIV-2 capsid sequence data was available from previous studies in the Caió cohort ([104], see Chapter 4 and appendix 1). HIV-1 subjects included in the study were selected on the basis of infection with HIV-1 CRF02_AG (determined by *pol* sequence analysis, see Chapters 2 and 3). HIV-1 and HIV-2 subtype was confirmed using the REGA HIV subtyping algorithm [232].

Sites under positive and negative selection in HIV-1 and HIV-2 capsid alignments were identified by comparison of synonymous and non-synonymous substitution rates using three different methods in the Datamonkey web-server [296]: SLAC, FEL and REL [297, 298]. Estimations were carried out using codon models of evolution selected by a genetic algorithm [299] on the Datamonkey suite and all alignments tested for recombination using the GARD tool [300]. A single recombination breakpoint in each of the HIV-1 and HIV-2 capsid alignments and therefore multiple partition datasets used for all analyses of selection. A p-value cut-off of 0.05 (SLAC and FEL) or Bayes Factor of 50 (REL) was used to define significant positive or negative selection at a codon.

6.3 RESULTS

Study subjects

Sixty HIV-2 mono-infected individuals were recruited into the study. Clinical and demographic parameters are described in Table 6.1. As expected, the proportion of HIV-2 infected individuals with over 13 years of follow up since HIV-2 diagnosis was higher in HIV-2 elite controllers (ECs). The predominantly female and older age of the study subjects is consistent with most West African HIV-2 cohorts [52, 336]. This may be due to an increased risk of HIV-2 acquisition in older women [337] and higher non-HIV-related mortality in men in sub-Saharan Africa [338]. The HIV-1 infected individuals included in the study are described in detail elsewhere [339] and summarised in Table 6.2.

Table 6.1. Characteristics of HIV-2 infected study subjects with undetectable (<100 copies/ml) and detectable plasma viral loads.

	Viral load < 100 (n = 33)	Viral load ≥ 100 (n = 27)	
Mean age in years (range)	62.3 (31 – 89)	62.7 (36 – 91)	NS
Female	79%	74%	NS
CD4 absolute count (cells/µl) Median (IQR)	596 (468 – 856)	400 (252 – 530)	P = 0.004
CD4% Median (IQR)	37.9 (30.7 – 43.9)	28.4 (19.8 – 38.2)	P = 0.005
Medial Viral Load copies/ml (IQR)	< 100	892 (693 – 20956)	NA
Follow up 13 – 21 years	27/33 (85%)	16/27 (59%)	P = 0.026

IQR = Inter-quartile range, NS = not significant, NA = not applicable

Table 6.2. Summary of HIV-1 infected slow and long-term non-progressors from whom samples were used in the study.

Subject	Year of HIV-1 diagnosis	Year of sample	CD4 at sampling (cells/mm^3)	Viral Load at samplings (copies/ml)	Tetramers used
NP005	1985	2002	539	108,789	A2-SLY (p24), A2-ILK (pol), B8-EIY (p24), B8-FLK (nef)
NP065	1985	2002	2310	18,817	B8-EIY (p24), B8-FLK (nef)
NP046	1985	2001	300	86,700	A2-SLY (p24), B8-EIY (p24), B8-FLK (nef)
NP200	1987	2000	405	497	A2-SLY (p24), B8-EIY (p24), B8-FLK (nef)

Control of HIV-2 viraemia is significantly associated with greater magnitude and polyfunctionality of the Gag-specific CD8+ T-cell response

Previous studies in the same cohort using IFN-γ ELISpots have shown a significant inverse correlation between the magnitude of Gag-specific IFN-γ responses from PBMCs and HIV-2 viraemia [172]. Further exploration in the current study (using intra-cellular cytokine staining, ICS) revealed that the CD8+ IFN-γ response, but not the CD4+ IFN-γ response, differentiated ECs from viraemic individuals ($p = 0.008$, Figure 6.1). A greater proportion of the Gag-specific CD8+ response consisted of IFN-γ and TNF-α producing cells in ECs (Figure 6.2a). The frequency of polyfunctional CD8+ T cells (i.e. those simultaneously producing ≥ 3 cytokines/soluble factors simultaneously in response to HIV-2 Gag) was significantly higher in ECs (Figure 6.2b) and polyfunctionality correlated inversely with plasma viral load (Figure 6.2c). In contrast, the proportion of the CD8+ T-cell response made up of mono-functional cells increased with higher VL (Figure 6.2d), providing further evidence that both the magnitude of HIV-2 Gag-specific CD8+ T-cells and the quality of the response are significantly associated with control of HIV-2 viraemia.

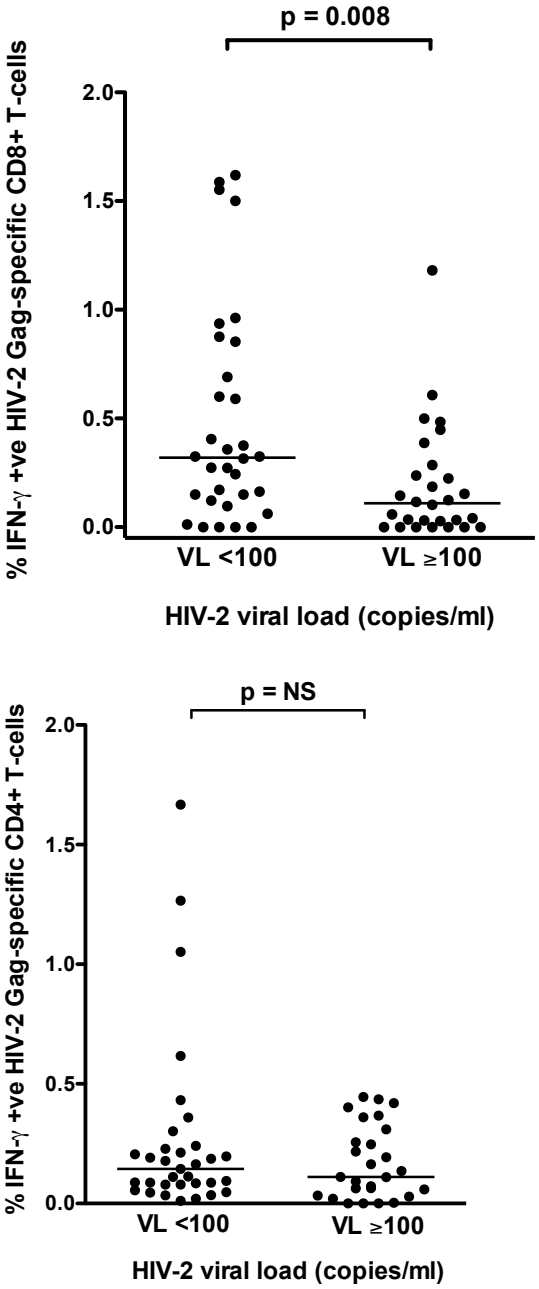


Figure 6.1. Comparison of HIV-2 Gag-specific CD8+ and CD4+ IFN-γ responses in HIV-2 infected individuals with undetectable and detectable plasma viral loads. Bars represent medians.

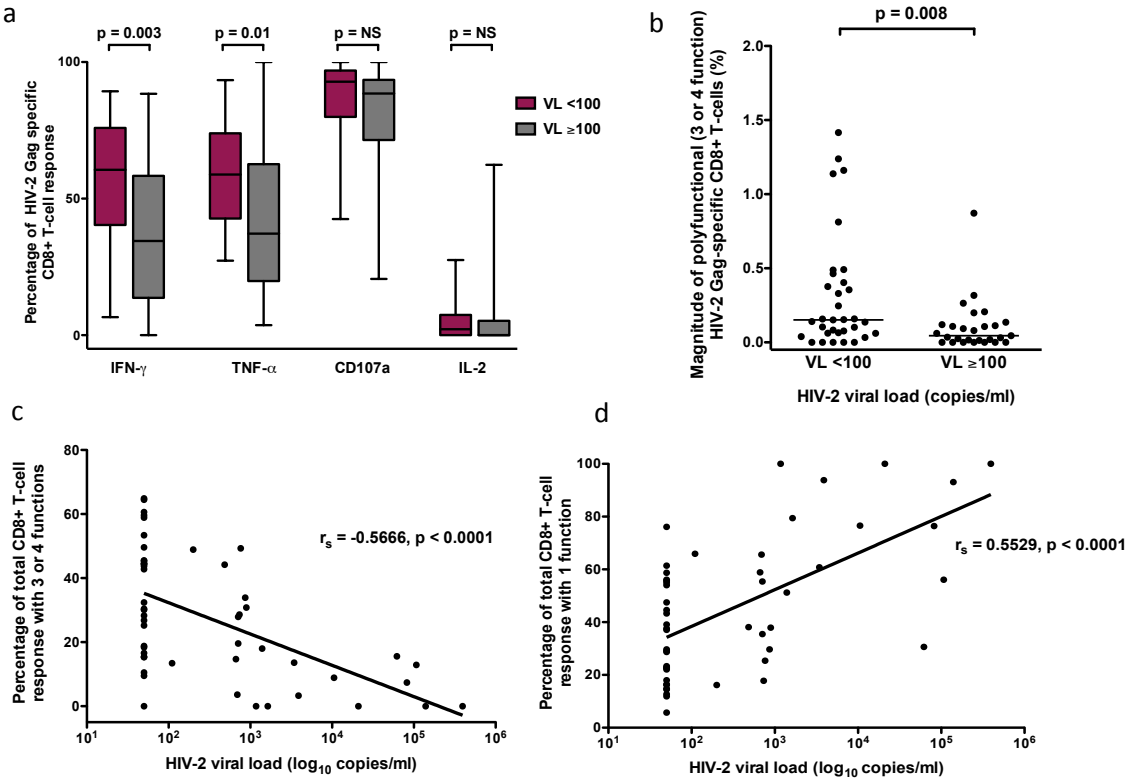


Figure 6.2. HIV-2 viral control is associated with polyfunctionality of the Gag-specific CD8+ T-cell response. (a) Proportion of response made up by different factors in elite controllers (ECs) and viraemic subjects (b) Magnitude of polyfunctional Gag-specific CD8+ T-cells in ECs and viraemic subjects (c) Inverse correlation between CD8+ polyfunctionality and HIV-2 viral load (d) Direct correlation between CD8+ monofunctionality and HIV-2 viral load. NS = not significant.

No significant association was observed between CD4+ count and the CD8+ IFN-γ response or polyfunctionality, demonstrating that the findings above are not simply a reflection of generalized immunosuppression with advancing HIV-2 disease progression (Figure 6.3). Finally, using median fluorescence intensity (MFI) of each parameter as a proxy for the amount of cytokine produced per cell, as previously described [170], polyfunctional Gag-specific CD8+ T-cells also secreted more IFN-γ (Figure 6.4) and TNF-α than their 'less functional' counterparts. A similar pattern for CD107a and IL-2 was not observed (data not shown).

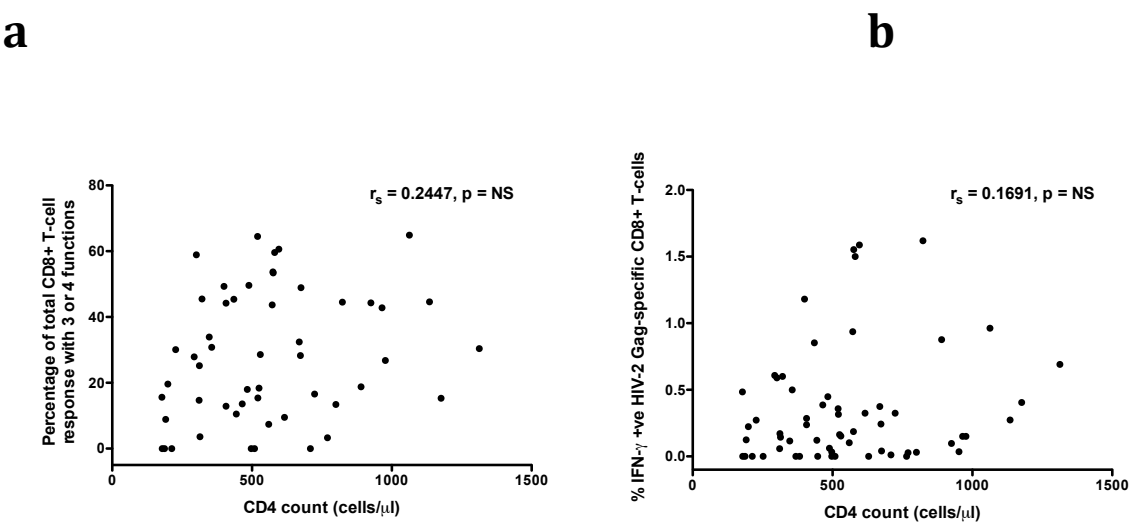


Figure 6.3. No association between CD4 count and (a) CD8+ polyfunctionality or (b) the magnitude of the HIV-2 Gag-specific CD8+ IFN-γ response. NS = not significant.

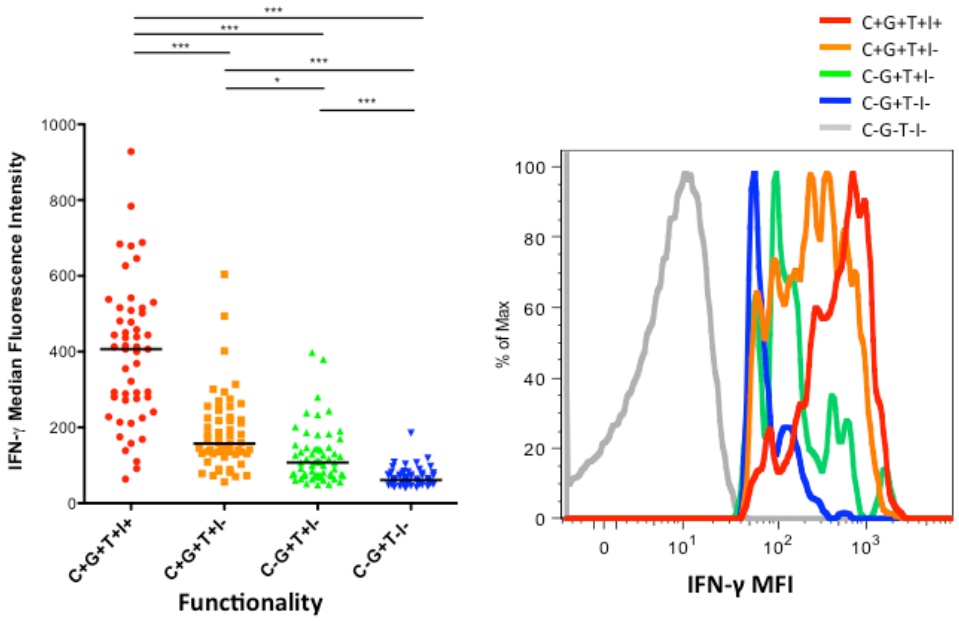


Figure 6.4. Relationship between Median Fluorescence Intensity (MFI) of IFN-γ and CD8+ polyfunctionality following HIV-2 Gag stimulation. (a) C,G,T and I represent CD107a, IFN-γ, TNF-α and IL-2 respectively. Comparisons made using Kruskal-Wallis test with Dunn's post-test for multiple comparisons. *** $p < 0.001$, * $p < 0.05$. (b) Representative histogram from a single subject.

The Gag-specific CD8+ T-cell response associated with HIV-2 elite control is highly focused and often absent in viraemic individuals

Peptide-MHC class I tetramers specific for four of the most immunodominant HIV-2 Gag epitopes described in this cohort [172] were used to determine the phenotype of HIV-2-specific CD8+ T-cells associated with control of HIV-2 viraemia. Thirty one of 56 individuals with available HLA typing (and sufficient PBMCs remaining) had B53, B58, B35 or B14 alleles (common HLA alleles in West Africa) and were included in the study (Appendix 2, a total of 35 tetramer staining experiments). Although phenotypes of

HIV-1, CMV and EBV-specific CD8+ T-cells are described in detail in the literature, 12 HIV-1-tetramers from four HIV-1 infected individuals (Table 6.2), eight CMV-specific tetramers (from five HIV-2 and three HIV-1 infected subjects) and 10 EBV-specific tetramers (from eight HIV-2 and two HIV-1 infected subjects) were also included to provide direct comparisons with HIV-2-specific CD8+ T-cells using the same flow cytometry panels.

A significantly greater proportion of HIV-2 ECs compared to viraemic subjects had tetramer-positive populations (84% vs. 42% respectively, $p = 0.02$, Figure 6.5b). Twenty five tetramer populations from 21 HIV-2 infected individuals (16 ECs, five viraemic) were therefore available for further characterization. A strong correlation ($r = 0.71$) was observed between the magnitude of the CD8+ IFN-γ response to the total HIV-2 Gag pool (by ICS) and the % of tetramer positive CD8+ T-cells in each subject (Figure 6.5c). This suggests that the bulk of the Gag-specific response in these subjects is accounted for by CD8+ T-cells targeting one or more of these four epitopes. A much stronger correlation was observed in ECs ($r = 0.8$, $p < 0.0001$) than viraemic subjects ($r = 0.62$, $p < 0.037$), which may represent a greater probability of Gag-specific CD8+ T-cells in viraemic subjects being functionally inert with reduced IFN-γ production, as has been described in a small proportion of HIV-1 progressors [340, 341].

To explore further the possibility that the Gag-specific CD8+ T-cell response associated with HIV-2 viral control is extremely narrow and focused, previously reported IFN-γ ELISpot data from the same cohort generated using a 3D-peptide matrix [172] was reanalysed. The total Gag pool-specific magnitude was significantly greater in individuals who responded to ≥1 of the above four HIV-2 Gag epitopes (40/64, 62.5%) than to those who did not (Figure 65d): median (IQR) 1103 (524 – 2081) SFU/ml vs.

195 (0 – 481) SFU/ml, thus confirming the remarkable immunodominance of these epitopes. Furthermore, of the 40 individuals who displayed responses to these B53, B58, B35 or B14 epitopes, 32 (80%) responded to only one of the four peptides. This represents an incredibly restricted breadth in the HIV-2-specific Gag response shown to be associated with viraemic control.

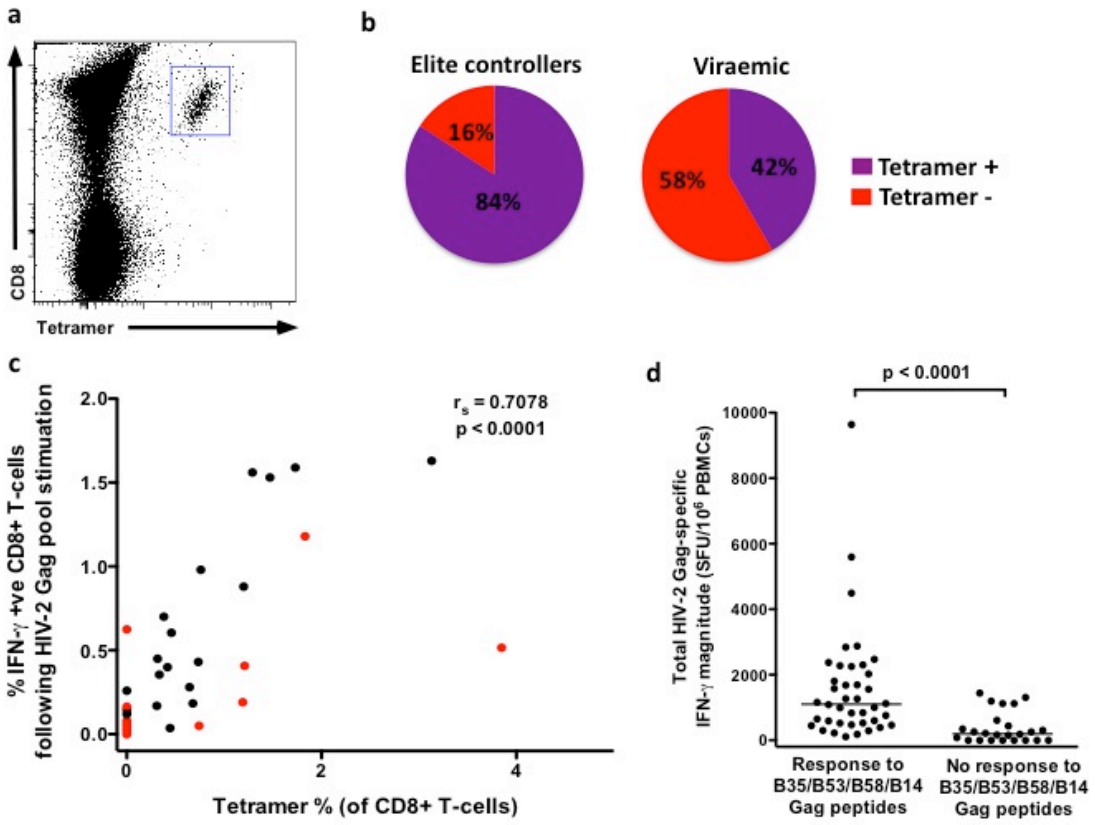


Figure 6.5. HIV-2 Gag-specific CD8+ T-cells enumerated by peptide-MHC class I tetramers reveal a highly focused response. (a) Example of tetramer staining (b) A greater proportion of HIV-2 elite controller have detectable immunodominant Gag-specific tetramer populations than viraemic individuals (p = 0.02) (c) Significant correlation between tetramer positive cells and IFN-γ response to the total Gag pool. Where 2 different tetramer responses were present

(in 4 subjects), these are added together. Red = viraemic donors (d) Total Gag-specific IFN-γ response according to presence or absence of immunodominant epitope response.

HIV-2-specific CD8+ T-cells display a distinct earlier differentiated phenotype from HIV-1 and CMV-specific CD8+ T-cells

Previous studies have shown significant heterogeneity in the phenotype of CD8+ T-cells specific for different viruses associated with chronic infection [342, 343]. HIV-2-specific CD8+ T-cells were significantly more likely to express CD27 than both HIV-1 and CMV-specific CD8+ T-cells, but not EBV-specific CD8+ T-cells (Figure 6.6a and 6.6c), representing T-cells at an earlier stage of differentiation. CD27 and CD45RO co-expression revealed a similar pattern. While CMV-specific T-cells were more likely to display a 'terminal effector' phenotype of CD27-/CD45RO-, HIV-1-specific T-cells were more frequently CD27-/CD45RO+ and HIV-2-specific T-cells CD27+/CD45RO+ (Figure 6.6b and 6.6c). No significant difference was seen in antigen-specific CD8+ T-cell senescence based on expression of CD57 (data not shown).

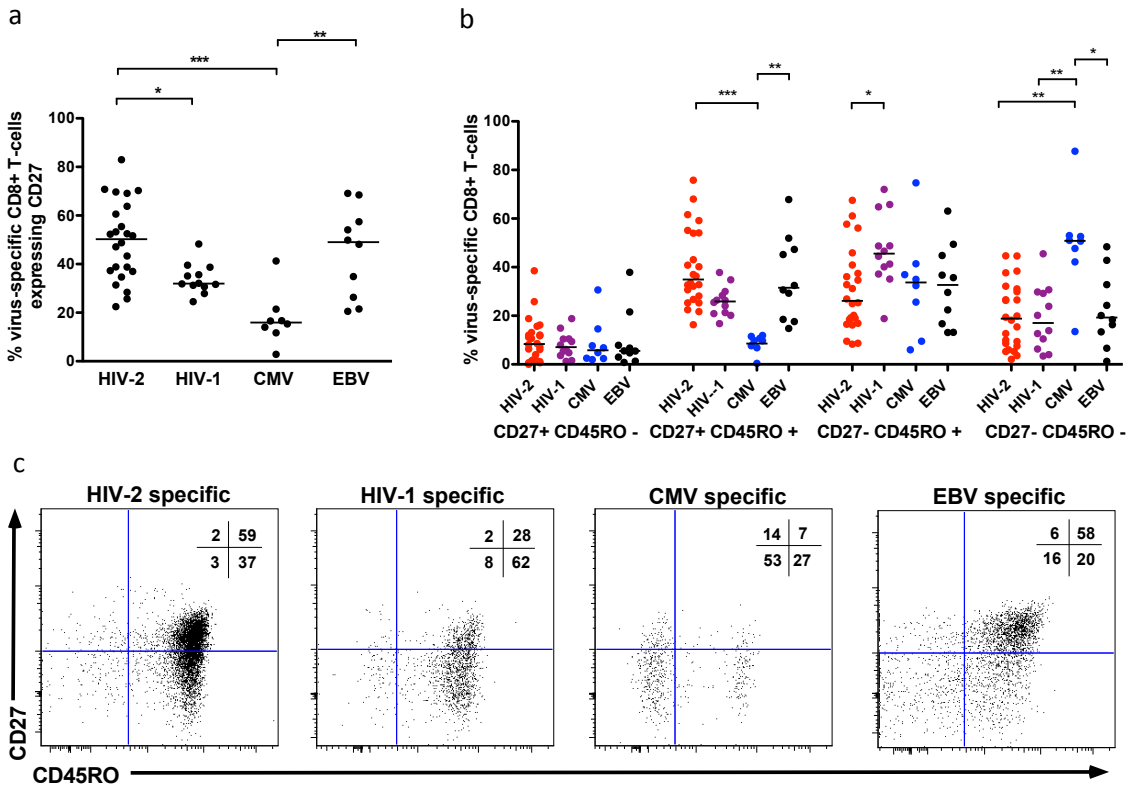


Figure 6.6. HIV-2 specific CD8+ T-cells display an earlier differentiated phenotype than HIV-1 or CMV specific CD8+ T-cells. (a) CD27 and (b) CD27/CD45RO co-expression according to viral specificity. (c) Examples of CD27/CD45RO distribution of HIV-2, HIV-1, CMV and EBV-specific CD8+ T-cells, with % distribution in each quadrant displayed. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$

HIV-2-specific CD8+ T-cells display high levels of PD-1, 2B4 and CD160 expression

Recent data have defined a relationship between co-expression of multiple inhibitory receptors (PD-1, 2B4, CD160) on HIV-1-specific CD8+ T-cells and higher viraemia, suggesting these represent a 'hyper-exhausted' cellular phenotype [344]. However, HIV-2-specific CD8+ T-cells (largely from individuals with undetectable plasma viral loads and polyfunctional Gag responses) have higher levels of PD-1, 2B4 and CD160 expression than bulk CD8+ populations (Figure 6.7a) and surprisingly similar levels of co-expression of these markers (Figure 6.7b) when compared to HIV-1-specific CD8+

T-cells. Furthermore, a significant direct correlation was observed between co-expression of PD-1, 2B4, CD160 and the magnitude of the IFN-γ response to Gag stimulation (r = 0.48, p = 0.027, Figure 6.8). No correlation was observed between polyfunctional responses to the Gag pool and co-expression of these markers on HIV-2-specific CD8+ T-cells (data not shown).

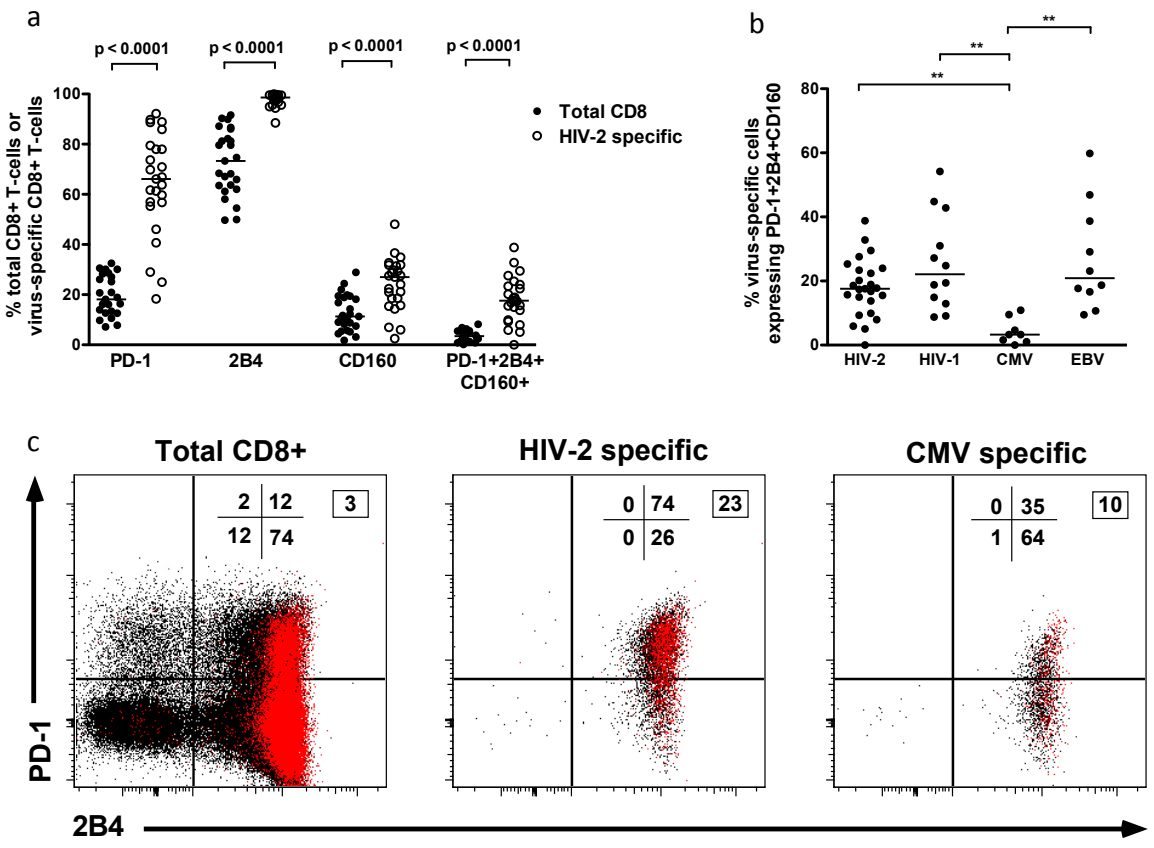


Figure 6.7. HIV-2 specific CD8+ T-cells show equivalent co-expression of PD-1, 2B4 and CD160 to HIV-1 and EBV specific CD8+ T-cells, but significantly lower than CMV specific CD8+ T-cells. (a) Expression of inhibitory receptors on bulk CD8+ and HIV-2 specific CD8+ T-cells (b) Comparison of PD-1+2B4+CD160+ co-expression on different virus-specific CD8+ T-cells. ** p < 0.01 (c) Example of inhibitory receptor expression on total CD8+, HIV-2 specific and CMV specific CD8+ T-cell from a single donor. Red dots represent cells co-expressing CD160 in

addition to PD-1 and/or 2B4. % distribution in each quadrant is displayed (PD-1 vs 2B4), with the boxed number denoting % of cells co-expressing all 3 markers.

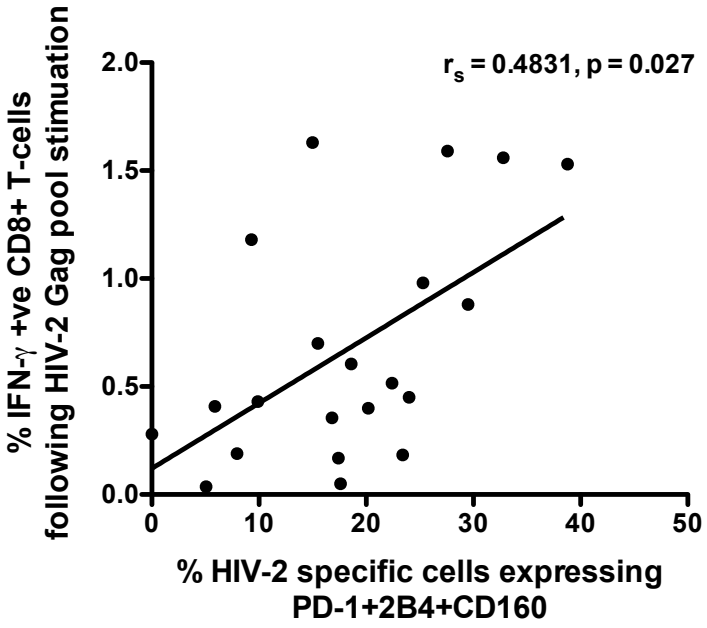


Figure 6.8. Correlation between co-expression of PD-1, 2B4 and CD160 and the magnitude of the HIV-2 Gag-specific CD8+ IFN-γ response.

CMV specific CD8+ T-cells displayed significantly lower co-expression of inhibitory receptors than all other virus-specific CD8+ T-cells (Figure 6.7b and 6.7c). It is possible that this observation may be related to the later differentiation phenotype of CMV specific T-cells. Accordingly, a strong correlation was observed between both PD-1 expression, as well as PD1, 2B4, CD160 co-expression, with CD27+CD45RO+CD8+ T-cells (Figure 6.9a and b) ($r = 0.69$, $p = 0.0002$ and $r = 0.75$, $p < 0.0001$ respectively). Thus expression of PD-1, 2B4 and CD160 is directly associated with the proportion of earlier differentiated CD8+ T-cells present.

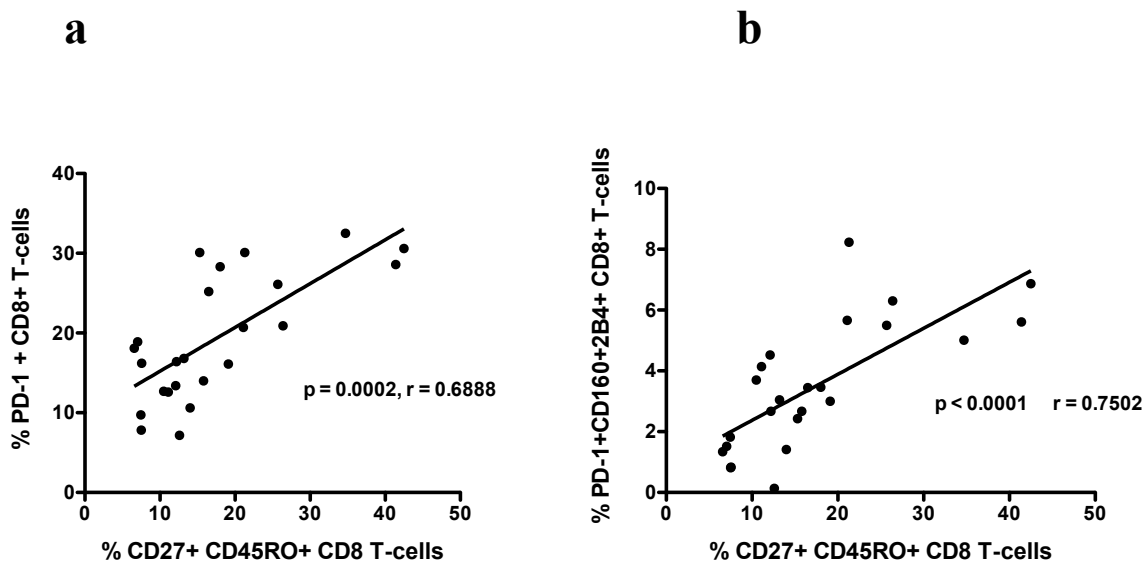


Figure 6.9. Correlation between (a) PD-1 and (b) PD-1+2B4+CD160+ expression and CD27+CD45RO+ CD8+ T-cells.

Control of HIV-2 viraemia is not associated with CD8+ T-cells with high cytotoxic potential

Control of HIV-1 viraemia has recently been shown to be associated with the capacity of HIV-1-specific CD8+ T-cells to express perforin on antigen stimulation, as well as the enrichment of other cytolytic markers such as Granzyme-B and T-bet [165, 345]. To explore whether similar qualities are required for HIV-2 control, the degree of perforin expression in response to Gag stimulation (defined as co-expression of perforin with one or more of IFN-γ, TNF-α, IL-2 and CD107a following stimulation [165]) was assessed. Although cells co-expressing perforin with other factors were readily identifiable following both SEB and Gag stimulation, there was no significant difference in this ability between HIV-2 ECs and viraemic subjects (Figure 6.10a and b). The degree of *ex vivo* expression of perforin, granzyme B and T-bet in virus-specific CD8+ T-cells using peptide-MHC class I tetramers was also assessed. When compared to bulk CD8+ T-cells, HIV-2-specific T-cells showed modest increases in granzyme B,

similar expression of T-bet and lower levels of perforin (Figure 6.10c). In keeping with their late differentiated phenotype, CMV specific CD8+ T-cells displayed significantly higher granzyme B and T-bet expression than HIV-2 and EBV specific T-cells, with perforin expression greater in HIV-1 when compared to HIV-2-specific T-cells (Figure 6.10d).

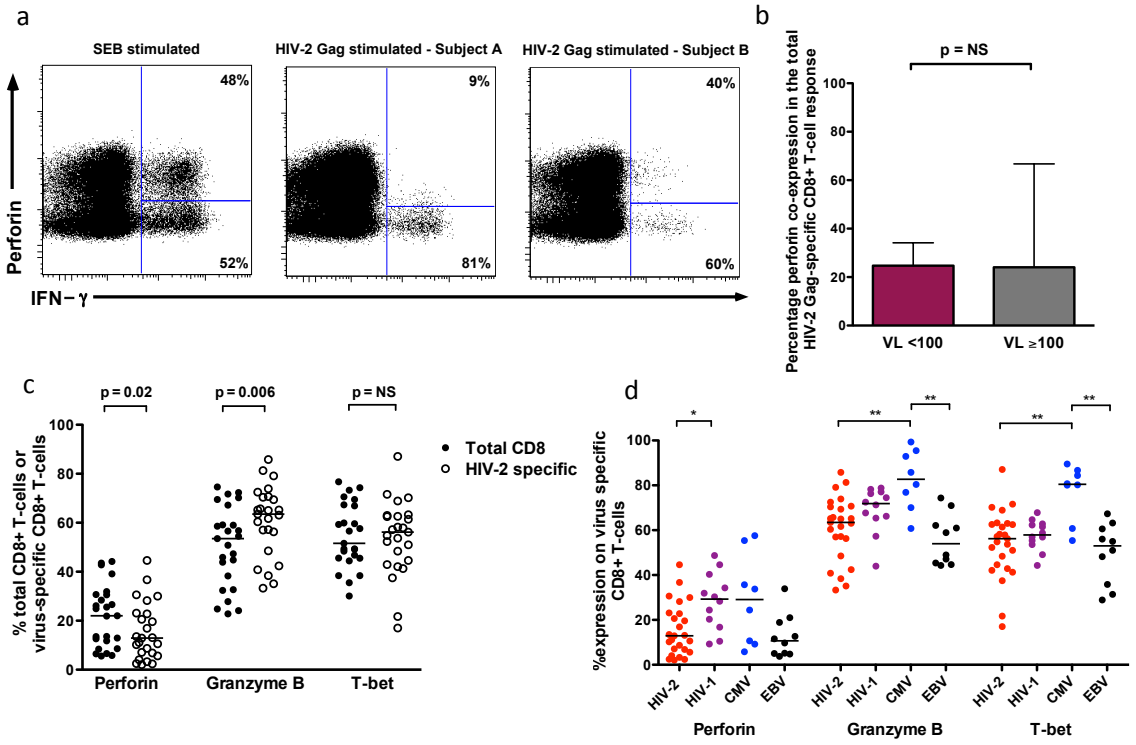


Figure 6.10. High cytotoxic potential is not a feature of HIV-2 specific CD8+ T-cells. (a) CD8+ T-cells co-expressing perforin and IFN-γ are detectable following both SEB stimulation, with varying proportions observed following HIV-2 Gag stimulation (b) No difference between HIV-2 elite controllers and viraemic subjects in the degree of perforin co-expression following Gag stimulation. Displayed are medians and upper IQR (c) Comparison of cytolytic markers on bulk CD8+ and HIV-2 specific CD8+ T-cells (d) Comparison of cytolytic markers on different virus-specific CD8+ T-cells. Bars represent medians. ** $p < 0.01$ * $p < 0.05$, NS = not significant.

HIV-2-specific CD8+ T-cells display low levels of activation and proliferation ex vivo.

Most virus-specific CD8+ T-cells display high levels of the activation markers CD38 and HLA-DR, as well as the cell cycle marker Ki67 during acute infection, which normalize rapidly in early chronic infection [342]. Generalized cellular immune activation has been shown to correlate with both viral load and disease progression in HIV-1 and HIV-2 [107, 109, 346], but the level of activation and proliferation of HIV-2-specific CD8+ T-cells is less well characterised. CD38+Ki67+ co-expression was significantly higher in HIV-1-specific CD8+ T-cells compared to both bulk CD8+ T-cells in the same individuals, as well as HIV-2-specific CD8+ T-cells (Figure 6.11a and b). In fact, when compared to bulk CD8+ populations, CD38, HLA-DR and Ki67 were all significantly elevated in HIV-1-specific but not HIV-2, CMV or EBV-specific CD8+ T-cells (data not shown). Interestingly, a significant inverse correlation was observed between CD38 expression on HIV-2-specific T-cells and the % CD8+ IFN-γ response following Gag pool stimulation (Figure 6.11c). A similar relationship was not observed between CD38 expression on bulk CD8+ T-cells and the IFN-γ response (Figure 6.12), suggesting activation of HIV-2-specific CD8+ T-cells *per se* may have a detrimental effect on function. Previous work has proposed an increased susceptibility to apoptosis in CD8+ T-cells expressing high levels of CD38 may result in a failure to control HIV-1 replication [347]. In keeping with this hypothesis, a strong inverse correlation was seen between CD38 expression and presence of Bcl-2, an anti-apoptotic marker, on HIV-2-specific CD8+ T-cells (Figure 6.11d).

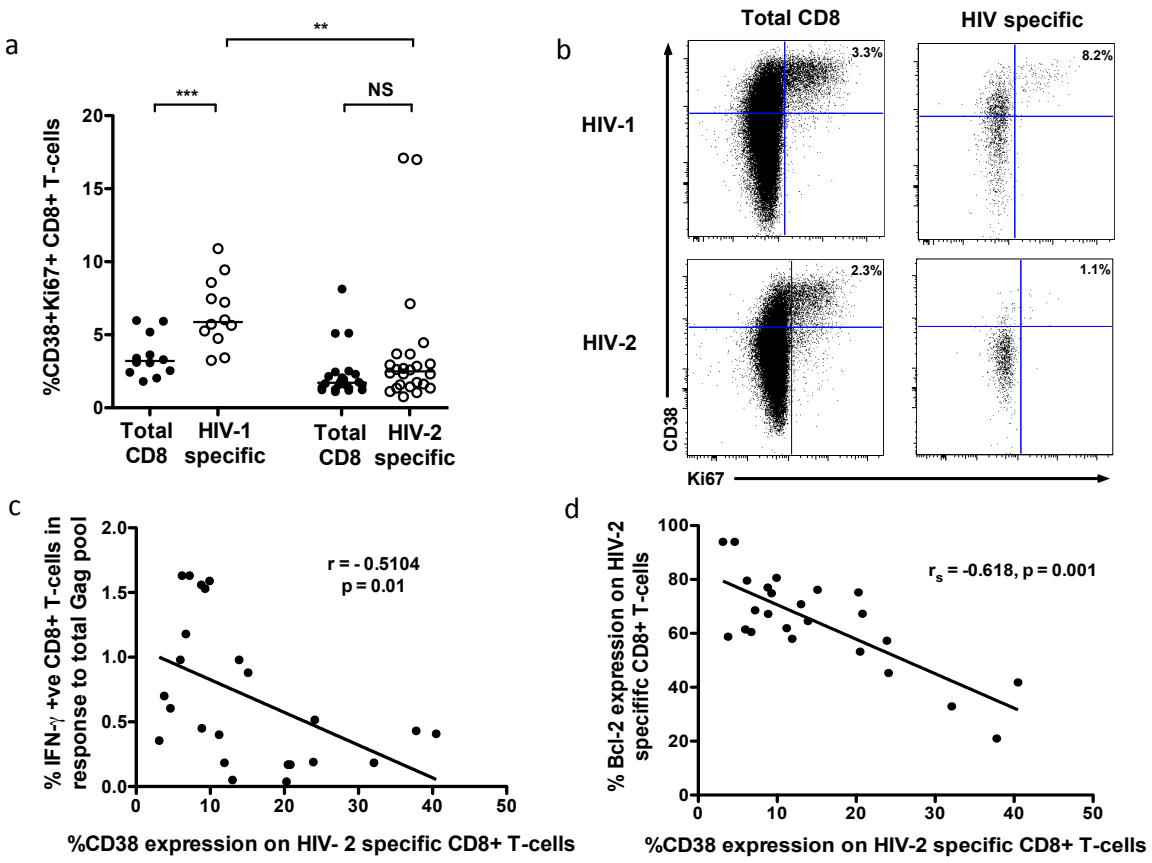


Figure 6.11. HIV-2 specific CD8+ T-cells display low activation and proliferation *ex vivo.* (a) Comparison of CD38+Ki67+ expression on HIV-1 specific, HIV-2 specific and bulk CD8+ T-cells. *** $p < 0.001$ ** $P < 0.01$, NS = not significant (b) Representative plots from an HIV-1 and HIV-2 infected donor showing CD38 and Ki67 expression on bulk CD8+ and HIV specific T-cells (c) Inverse correlation between activation of HIV-2 specific CD8+ T-cells and CD8+ IFN-γ response to the Gag peptide pool (d) Inverse correlation between activation and the expression of the anti-apoptotic marker Bcl-2 on HIV-2 specific CD8+ T-cells.

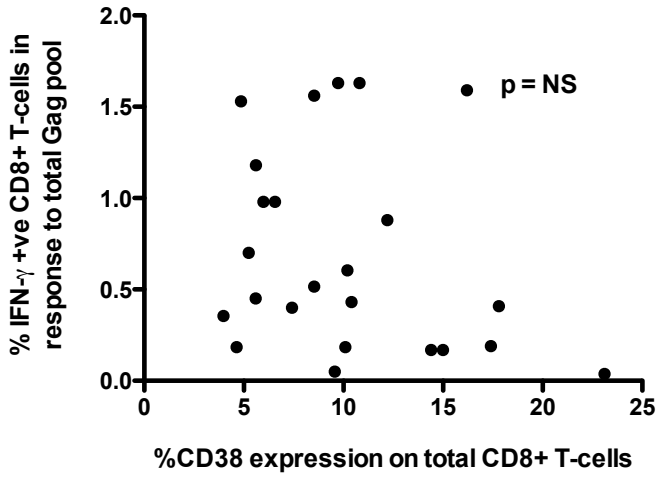


Figure 6.12. No correlation between expression of the activation marker CD38 on the total CD8+ T-cell population and the IFN-γ response to stimulation with HIV-2 Gag.

The HIV-2 capsid is under greater negative selective pressure than the HIV-1 capsid in the same community cohort

Although due to functional constraints the HIV-1 capsid is an extremely conserved region of the HIV-1 genome, immune pressure from HIV-1-specific T-cells can result in amino acid mutations within the epitopes targeted and as a consequence lead to immune escape (reviewed in [327, 348]). Given the potent CD8+ T-cell responses observed in HIV-2 infected Caió subjects, the selective pressures on the HIV-2 and HIV-1 capsids were compared in the same community. HIV-2 capsid sequences from 86 Caió subjects were available from previous studies ([104], see Chapter 4 and appendix 1). HIV-1 capsid sequences from 55 Caió subjects known to be infected with HIV-1 CRF02_AG (via *pol* sequence analysis, see Chapter 3) were obtained. The number of sites under negative selective pressure was significantly greater in the HIV-2 capsid compared to the HIV-1 capsid using both the SLAC and FEL algorithms (Table 6.3, Figure 6.13). Accordingly, the mean dN/dS

score (non-synonymous/synonymous mutations) was significantly lower in the HIV-2 capsid (Table 6.3). Of the six sites under significant positive selective pressure in the HIV-1 capsid, three were within known epitopes restricted by HLA alleles common in Caió (Figure 6.13). In contrast, the four HIV-2 immunodominant epitopes described above did not display any sites significantly altered due to positive selective pressure in Caió. Thus the functional constraints present which prevent amino acid mutations in the HIV-2 capsid may be far greater than in the HIV-1 capsid., preventing immune escape from T-cells.

Table 6.3. Selective pressure on HIV-1 (n = 55) and HIV-2 (n = 86) capsid sequences in Caió.

Selection analysis @	**HIV-1***	**HIV-2***	***P* value**
SLAC			
No. of positively selected sites	6	2	< 0.0001
No of negatively selected sites	61	139	
Mean dN/dS (95% confidence intervals)	0.25 (0.22 - 0.28)	0.1(0.09 - 0.11)	< 0.0001
FEL			
No. of positively selected sites	8	2	< 0.0001
No of negatively selected sites	82	157	
REL		NA**	
No. of positively selected sites	13	NA	
No of negatively selected sites	85	NA	

@ Sites considered under significant positive or negative selection if $p < 0.05$ (SLAC and FEL) or Bayes factor >50 (REL).
* HIV-1 = 231 codons, HIV-2 = 230 codons
** Not performed due to alignment size restrictions

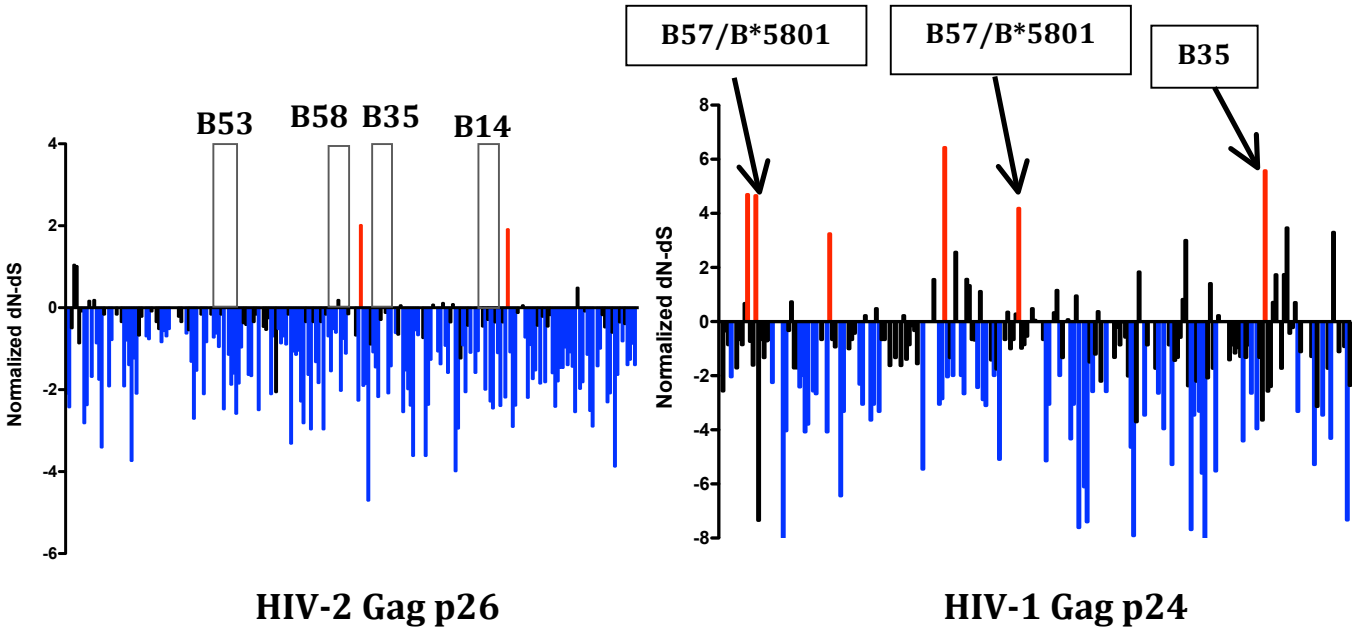
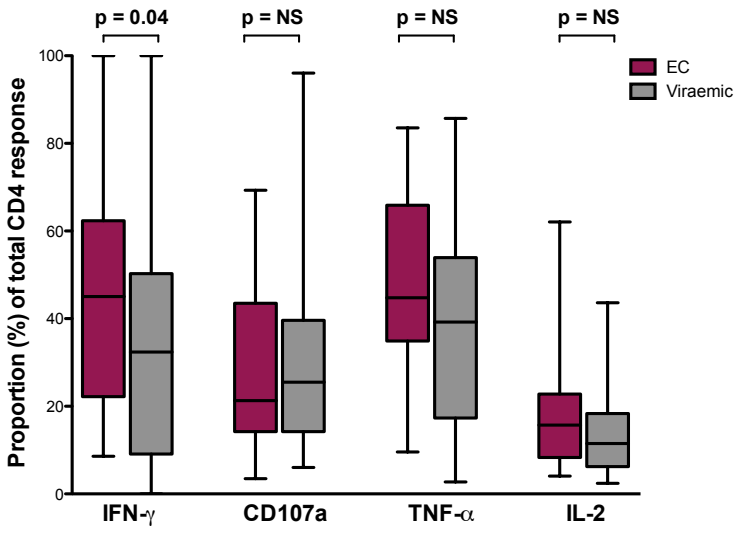


Figure 6.13. Comparison of selective pressure on HIV-2 (n = 86) and HIV-1 CRF02_AG (n = 55) capsid sequences. Plotted values are those obtained under the SLAC algorithm with values normalized to alignment length. Positions found to be under significant positive selection (red) and negative selection (blue) under both SLAC and FEL algorithms are highlighted. For HIV-2 Gag p26, the position of the four immunodominant epitopes (restricted by HLA B53, B58, B35 and B14) are shown. For HIV-1 Gag p24, sites under positive selective pressure within known T-cell targeted epitopes are labeled, along with the respective HLA restriction.

The composition of the HIV-2 Gag specific CD4+ T-cell response

Despite the magnitude of the CD4+ HIV-2 Gag-specific response (as assessed by IFN-γ secretion) not being greater in ECs than viraemic individuals, it is possible that the composition (i.e. proportion of Gag-specific cells producing a variety of cytokines) may distinguish the two groups. IFN-γ secreting cells formed a significantly higher proportion of the CD4+ Gag-specific response in ECs (p = 0.04, figure 6.14a), with no differences seen in the proportion of CD107a, TNF-α or IL-2 expression. ECs have a more polyfunctional response, with triple and quadruple cytokine-secreting cells making up a greater proportion of the response (figure 6.14b). The magnitude, however, of these two populations was not significantly different in the two groups (p = 0.17 and p = 0.06 respectively).

A



B

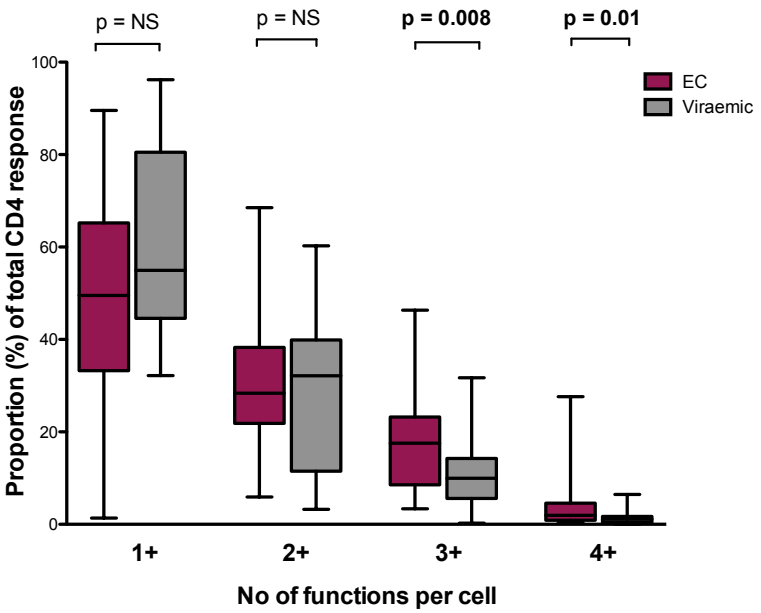


Figure 6.14 (a) Proportion of the total Gag-specific CD4+ T-cell response contributed by each function and (b) The proportion of the total Gag-specific CD4+ T-cell response contributed by cells with different degrees of functional ability. Shown are boxes representing medians and inter-quartile ranges, with whiskers representing range, for elite controllers (EC) and viraemic individuals. The magnitude of the total CD4+ Gag-specific response was calculated by summation of all possible response patterns for the 4 functions described. Differences were assessed by the Mann-Whitney U test. NS = not significant.

6.4 DISCUSSION

A key factor in designing an effective T-cell vaccine for HIV-1 infection is the clear identification of immune correlates associated with viraemic control. The strong class I HLA associations [122] and the association of a broad gag-specific T-cell response with lower viral loads [349] implicate CD8+ T-cells as a key element of HIV-1 control. Furthermore, the successful control of pathogenic SIVmac239 infection in rhesus

macaques via induction of effector memory T-cell responses [152] provide cause for optimism. Nevertheless, the failure of the Merck Ad5-HIV vaccine [350] demonstrates that the goal of an effective anti-HIV-1 T-cell vaccine in humans remains elusive. HIV-2 infection provides an alternative model of natural HIV control, yielding important insights into HIV pathogenesis and viral control in humans. The current study provides detailed information about the qualities of CD8+ T-cells associated with natural HIV-2 control. Polyfunctionality, but not high cytotoxic potential, is associated with viraemic control and accordingly, HIV-2-specific CD8+ T-cells display an earlier differentiated phenotype with low levels of activation and proliferation *ex vivo*. Additionally, these data show that a very focused and narrow but highly functional Gag-specific response is associated with excellent control of HIV-2 viraemia.

Using IFN-γ ELISpots, previous studies in the same cohort have highlighted the immunodominance of Gag in HIV-2-specific T-cell responses associated with undetectable HIV-2 VL [172]. The current studies demonstrates that this correlation is largely due to the CD8+ response, although the contribution of CD4+ help by secretion of cytokines such as IL-21 was not assessed and may play a role [351]. The more polyfunctional CD4+ responses seen in ECs when compared to viraemic individuals suggest that the quality of the CD4+ response may well be important. The role of broad, immunodominant Gag-specific CD8+ responses in HIV-1 control has also been shown, but requires large studies to demonstrate a significant association [349, 352]. The hierarchical dominance of Gag-specific responses in HIV-2 is, however, more profound [172], whereas Nef-specific responses can dominate in HIV-1 infection and have also been associated with viral control [353, 354]. Maintaining a substantial breadth of T-cell responses is considered important in HIV-1 control and is associated with greater polyfunctionality and virus inhibition [355]. The requirement for broad host responses,

as a result of HIV-1 diversity and adaptation, poses a significant challenge for HIV-1 vaccine design. The current data suggest a stark contrast in HIV-2 infection, where CD8+ T-cells responding to a single HIV-2 epitope are responsible for almost the entire Gag-specific response in a particular subject. Furthermore, the absence of tetramer populations in most viraemic HIV-2 subjects marks another difference with HIV-1 infection, where the persistence of high levels of tetramer-staining CD8+ T-cells is frequently observed despite high viraemia and progression to AIDS [356]. Long-term control of HIV-2 viraemia, therefore, appears to be critically dependent on the presence of a very focused, high avidity [174] CD8+ response to conserved p26 epitopes, which is mostly functional when present. Interestingly, the HLA-B27 restricted KK10 response is an example of a single immunodominant high avidity p24 response responsible for controlling HIV-1 viraemia *in vivo* [175]. Along with the data presented in this chapter, this supports the hypothesis that in the presence of a single highly functional immune response, increasing the breadth of anti-viral specificities may not be necessary [357].

It is clear that different chronic viral infections are characterised by virus specific CD8+ T-cells displaying a particular cellular phenotype [342], and, accordingly, their effector functions may also vary. Several studies demonstrate that rare individuals who control HIV-1 replication in the absence of ART possess HIV-1-specific CD8+ T-cells that are more differentiated than viraemic subjects, with downregulation of markers such as CD27 and CD45RO [159, 345, 358], whereas other studies have shown that the phenotype of tetramer-staining cells does not differ significantly between controllers and progressors [359]. The more differentiated CD8+ T-cell subsets described in some HIV-1 ECs also have the ability to rapidly upregulate cytotoxic granule contents such as perforin and granzyme B on stimulation, and express higher levels of the transcription

factor T-bet [164, 165, 345], suggesting that HIV-1 control may require potent cytotoxic potential. Surprisingly, HIV-2-specific CD8+ T-cells (largely from individuals with low or undetectable VL) are characterised by low levels of pre-loaded cytotoxic molecules, and that rapid upregulation of perforin after Gag stimulation does not distinguish HIV-2 ECs from viraemic subjects. In keeping with this observation, HIV-2-specific CD8+ T-cells are of an earlier differentiated phenotype, with greater expression of CD27 and CD45RO. Polyfunctionality is, however, important in HIV-2 control and is a feature shared with HIV-1-specific CD8+ T-cells from HIV-1 controllers [161, 175].

The reasons for the earlier differentiation status of HIV-2-specific CD8+ T-cells are intriguing. Studies exploring the relationship between antigen burden and cellular differentiation offer a potential explanation [173, 360]. Using both *ex vivo* antigen-specific CD8+ T-cell analysis and a model of *in vitro* T-cell priming, a previous study suggests that higher viral burden drives T-cell differentiation in HIV infection through greater immune activation [173]. Consistent with these data, HIV-1 subtype C infected individuals with higher VL set points following acute infection have greater frequencies of differentiated effector memory HIV-1-specific CD8+ T-cells (that are CD27-), along with higher activation (CD38+) levels [360]. Although studies of HIV-2 acute infection are lacking, it is likely that the VL set point of HIV-2 ECs is low or undetectable at an early stage which may explain the persistent CD27 expression observed in chronic infection. The low activation levels observed in HIV-2-specific CD8+ T-cells in the current study are also in keeping with this hypothesis.

Alternatively, the distinct differentiation phenotypes observed in different chronic viral infections may be related to the distinct qualities required by CD8+ T-cells to control different viral infections. This could explain the finding that HIV-2-specific CD8+ T-

cells from ECs are defined by polyfunctionality but not cytotoxicity, whereas increasing evidence suggests direct contact with infected CD4+ T-cells and potent cytotoxicity is required to control HIV-1 replication [159, 164, 165, 345]. It is therefore possible that the inhibitory mechanisms required to contain HIV-1 and HIV-2 *in vivo* are not identical. Interestingly, cell supernatants from HIV-2-specific CD8+ T-cell clone cultures are able partially to suppress HIV-2 replication in CD4+ T-cell lines without cell-to-cell contact (Aleksandra Leligdowicz, unpublished data). High co-expression of PD-1, CD160 and 2B4 is also evident on HIV-2-specific CD8+ T-cells. Although robust data exist to suggest an inhibitory role for these markers in HIV-1 disease [344], the current findings support previous work suggesting PD-1 upregulation may also reflect changes in cellular differentiation [361]. In addition, it is possible that expression of these markers has less effect on cytokine-mediated control of HIV-2, but are more relevant in inhibiting cellular mechanisms required for efficient HIV-1 viral suppression.

The current study is relatively limited by being based in a single cohort from Guinea-Bissau, although the length of follow up in a genetically homogeneous population from the same community infected with the same HIV-2 group (A) is also a major strength that is difficult to reproduce in HIV-1 controller cohorts. It will be important to corroborate these findings in other HIV-2 cohorts, especially to see if a focused high avidity Gag-specific response associated with HIV-2 control is common to subjects with different HLA allele distributions. How an HIV-2-specific CD8+ response with a very narrow specificity is able to control HIV-2 viraemia for decades warrants further exploration. Such responses with high functionality but low breadth remain beneficial only if viral escape within the restricted epitope does not occur. This would result in uncontrolled viral replication in the absence of CD8+ T-cell responses to other regions, which has been observed in the case of the HLA B27-

restricted response to KK10 in HIV-1 [156]. Despite the fitness costs associated with mutation, escape from Gag-specific CD8+ T-cell responses are frequently seen in HIV-1 infection (reviewed in [327, 348]). Selection analyses comparing HIV-1 and HIV-2 *env* [268] and *gag* reveal that HIV-2 is under greater negative selective pressure than HIV-1, suggesting far greater functional and structural constraints are imposed on HIV-2 compared with HIV-1. Escape from CD8+ T-cell responses in HIV-2 has not been described and autologous viral sequences from individuals with potent Gag-specific IFN-γ ELISpot responses show a high degree of conservation in respective epitopes, despite the presence of strong immune selective pressure [174]. It is possible that due to greater constraints against changes in the HIV-2 capsid, the first CD8+ T-cell responses that are mounted against conserved epitopes can persist and continue to control viraemia for years.

Another key limitation worth discussing is common to most studies exploring immune correlates of protection in natural HIV infection, namely, the difficulty in determining whether the potent anti-HIV-2 CD8+ T-cell responses observed represent the *cause* of HIV-2 viral control, or if they are simply a reflection of an intact immune system due to limited viral replication secondary to other factors (either viral or host immune). Several observations argue for a causative role. Remarkably high neutralising antibody responses in HIV-2 infection are evident (Chapter 5 and [310, 362]), but in contrast to T-cell responses, there is no association between the magnitude or breadth of this response and viral control [268]. The current analyses also fail to show an association between CD4+ T-cell counts (a reliable marker of immunodeficiency) and the magnitude or quality of the HIV-2 Gag specific CD8+ T-cell response.

Nevertheless, the strongest evidence for a causative role would be established by achieving elite control of HIV-2 in viraemic individuals via immuno-therapeutic strategies designed to induce CD8+ T-cell responses with the characteristics described in this study. The current data should, therefore, pave the way for developing CD8+ T-cell based therapeutic strategies against HIV-2. An HIV-2 Gag-based therapeutic T-cell vaccine could provide proof-of-concept for future induction of HIV-1 immune-mediated control and allow valuable lessons to be learned about robust vaccine-induced control of HIV infection. This would represent a vital step closer to an HIV-1 vaccine from the recent success of anti-SIV vaccines in animal models [152, 363]. In addition, it may provide a non-pharmacological therapeutic option for clinical care of HIV-2 infected subjects, who are often disadvantaged by sub-optimal ART regimes [6], compounded by the lack of reliable drug procurement in resource-limited settings.

CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS

Despite 30 years passing since the beginning of the HIV-1 pandemic, HIV pathogenesis is still not fully understood. Combination ART for HIV-1 has been a tremendous success and HIV-related mortality has fallen precipitously in countries with good access to ART. Continued advances in this field provide much reason for optimism. Nevertheless, the cumulative toxicity of long-term ART is an increasing problem and both therapeutic and prophylactic HIV-1 vaccines are urgently required. It is clear that further insight into the interplay between HIV and host immunity is needed to achieve these goals, in a way that was not required in the era of empirical vaccinology. Exploring models of naturally acquired immunity to HIV is a key strategy in gaining this knowledge. The majority of this thesis has been dedicated to the study of HIV-2, where most infected individuals remain asymptomatic, but approximately 15 – 20% progress to AIDS akin to HIV-1 infected subjects. Establishing the reasons for the attenuated disease course in most HIV-2 infected subjects infected with a potentially pathogenic retrovirus could help develop novel approaches to controlling HIV-1.

Many previous studies have focused solely on comparing differences between HIV-1 and HIV-2 infected individuals, often showing no differences between the two. Yet it is clear that HIV-2 infected individuals should not be treated as a homogenous group who all progress to disease more slowly than their HIV-1 infected counterparts. Studies in Caió have revealed that some subjects maintain their elite controller LTNP phenotype for over a decade, whereas others are viraemic and progress to AIDS at variable rates. What is clear, is that HIV-2 VL, along with generalized immune activation, are strong predictors of this difference

in HIV-2 outcome. If insight is to be gained into HIV pathogenesis via the study of HIV-2, it is vital that future studies address both of these questions: (1) what differentiates the course of HIV-1 from HIV-2 infection and (2) what determines whether an individual infected with the same strain of HIV-2 behaves like an HIV uninfected individual or an HIV-1 infected subject? It is likely that the answers are multifactorial and lie within both viral and host factors.

It is possible that more or less virulent strains of HIV-2 exist, with elite controllers infected with the latter and intrinsic viral factors may therefore dictate the divergent courses of HIV-2 infections. However, data presented in Chapter 4 clearly demonstrate that HIV-2 infected individuals whose viruses share a most recent common ancestor (and in some cases are known sexual partners) can be either viraemic subjects or elite controllers. This argues against a strong viral determinant of pathogenicity and is an important finding that suggests host factors play a key role in determining the outcome of HIV-2 infection. It is still possible that certain viral motifs can be identified, that are enriched in elite controllers, which may of course reflect adaptation secondary to host immune pressure. An example of this is the association between the presence of three prolines at positions 119, 159 and 178 in the HIV-2 capsid and undetectable plasma VL [104]. Phylogenetic and ancestral state studies show that these variants occur throughout the diversity of HIV-2 isolates and are prone to selection or reversion during transmission ([104] and Matt Cotten, unpublished data). Future studies should focus on determining what causes selection of these prolines and examining the whole HIV-2 genome to identify further motifs associated with elite control of HIV-2. It is worth noting that most of these studies to date have been performed on HIV-2 group A isolates and very little is known about HIV-2 group B,

and even less about the recently described HIV-2 recombinant, CRF01_AB. Epidemiological data on HIV-2 infections from many parts of West Africa are lacking and it is also worth establishing whether the transmission and natural history of HIV-2 is similar in these settings to what has been described in cohorts from the Gambia, Guinea-Bissau and Senegal.

The recent discovery that SIVcpz in its natural host, the chimpanzee, can in fact give rise to immunosuppression and AIDS [364] adds a new dimension to the understanding of HIV pathogenesis. It was previously assumed that SIVcpz, like SIVsmm in sooty mangabeys, was asymptomatic in chimpanzees. Why is it that during the cross-species transmission of SIVcpz to HIV-1 many consequences for the host have been retained, whereas HIV-2 in humans behaves so differently from SIVsmm in sooty mangabeys? Furthermore, it is intriguing that zoonotic transmission from two different simian viruses have resulted in identical clinical features of HIV-1 and HIV-2 in some humans. The role of different viral proteins in antagonising the host restriction factor tetherin deserves consideration in this context. SIVcpz Nef is able to counteract chimpanzee tetherin, whereas HIV-1 group M Vpu has adopted this ability against human tetherin, presumably as human tetherin is resistant to degradation by SIVcpz and HIV-1 group M Nef. Interestingly, amongst the different HIV-1 groups described so far, there is only evidence for efficient tetherin antagonism by HIV-1 group M Vpu. It is possible, therefore, that this quality (by aiding transmission) is partially responsible for the pandemic nature of this particular HIV-1 group. Determining whether HIV-2 Env's ability to antagonise tetherin is less efficient (than HIV-1 group M Vpu) and whether Env proteins from viraemic subjects and elite controllers vary in this function may aid the understanding of HIV pathogenesis.

More recently attention has focused on HIV-2 Vpx, which has the ability to counteract SAMHD1, the HIV-1 restriction factor found in cells of the myeloid lineage. While much attention has been dedicated to the role of SAMHD1 in HIV-1 pathogenesis, it is also important to consider whether this functional difference between HIV-1 and HIV-2 can explain, in part, the markedly different disease outcomes in infected subjects. Why might a greater ability to infect myeloid lineage cells lead to enhanced viral control in HIV-2? One possibility is that infection of cells such as dendritic cells results in greater levels of antigen presentation and therefore superior host immune responses than those observed in HIV-1 infected subjects. Future work should try to establish whether there is variability in the ability to counteract SAMHD1 between HIV-2 Vpx alleles derived from different subjects and if this is associated with VL and/or anti-HIV-2 immune responses.

If host immune responses rather than intrinsic viral properties are responsible for dictating the varying outcome of HIV-2 infections, further characterisation of this potentially protective response is important. Data presented in Chapter 5 arguably represents the most comprehensive exploration of neutralising antibody responses (Nabs) in HIV-2 infection to date. Remarkably high titres of Nabs are found in most HIV-2 infected subjects and are several fold higher than seen in HIV-1 infection. Yet no association between the magnitude (or breadth) of these heterologous or autologous Nab responses and better viral control is observed. Paradoxically, the magnitude of the heterologous response is directly correlated with plasma VL, suggesting antigen burden driven B-cell activation is common to both HIV-1 and HIV-2 infections. Despite the discovery of several HIV-1 infected 'elite neutralisers' from whom potent monoclonal antibodies have been isolated,

there is also no direct evidence that Nabs in HIV-1 help control VL *in vivo*. Whether the high Nab levels still contribute to slower progression in HIV-2 infected subjects (although they are not the main determinate) when compared to HIV-1 infected subjects, is not clear. In addition, why are Nab titres in HIV-2 infected subjects so much higher than those seen in HIV-1 infected individuals? B-cell function could be better preserved in HIV-2 infection or alternatively, the HIV-2 envelope may be uniquely susceptible to neutralisation. Demonstration that B-cell depletion is similar in HIV-1 and HIV-2 [146], as well as structural modeling suggesting the HIV-2 envelope is more accessible to Nab [310], indicates that differences between the biological properties of HIV-1 and HIV-2 envelopes are responsible.

Archived plasma samples from Caió subjects from up to 15 years previously were able to neutralise autologous envelopes isolated from recently sampled time points. This indicates that highly neutralisation sensitive envelopes are allowed to persist in the circulation for years despite antibody pressure – a stark contrast to HIV-1 where envelope escape from Nab occurs rapidly in early infection. These findings are supported by selective pressure analyses from Caió demonstrating that the HIV-2 envelope is under significantly greater negative selective pressure than the HIV-1 envelope. It is possible, therefore, that the HIV-2 envelope is under certain functional constraints, that prevents escape from Nab in most infected subjects. Yet neutralisation resistant envelopes are found in some subjects and the mechanisms of escape (e.g. changes in N-linked glycosylation sites) may resemble what is seen in HIV-1. These resistant envelopes are found more often in HIV-2 subjects with higher viraemia and greater intra-patient viral diversity. It is possible, therefore, that in the majority of HIV-2 infected subjects, host factors act

to limit HIV-2 replication and diversification to an extent that escape from Nab is prevented.

Could it be cellular immune responses that prevent HIV-2 replication to such an extent? As CD8+ T-cell responses appear early in the course of HIV-1 infection and play a role in controlling in peak viraemia, it is plausible that potent T-cell responses are key in determining outcomes in HIV-2 infected subjects. The data presented in Chapter 6 builds on prior work in the Caió cohort to support this hypothesis. Using intra-cellular cytokine staining, the magnitude of the CD8+ IFN-γ response following Gag stimulation differentiated HIV-2 elite controllers from viraemic subjects. Furthermore, this response was more polyfunctional in elite controllers, whereas viraemic individuals had predominantly monofunctional Gag-specific responses. Using peptide-MHC class 1 tetramers to identify HIV-2 specific CD8+ T-cells targeting 4 immunodominant Gag epitopes, further insight into cellular immunity to HIV-2 infection was gained. HIV-2 elite controllers were more likely to have circulating Gag-specific CD8+ T-cells identified than viraemic subjects. This is in contrast to HIV-1 infection where tetramer-positive populations are often identified in individuals with high VLs and advanced disease; although these populations are not always functional. HIV-2 Gag-specific CD8+ T-cells are also at an earlier state of differentiation than HIV-1 infected CD8+ T-cells and accordingly are not highly activated or enriched for cytolytic markers. These findings would be in keeping with a model where lower VLs in early HIV-2 infection result in lower activation and differentiation of Gag-specific CD8+ T-cells. Alternatively, this specific phenotype of HIV-specific CD8+ T-cells (polyfunctional but not highly cytotoxic) may be best suited to control HIV-2, whereas HIV-1

control requires a more differentiated and cytotoxic CD8+ T-cell response, as demonstrated through recent studies of HIV-1 elite controllers.

Another striking finding was that the HIV-2 Gag-specific CD8+ T-cell response is extremely narrow, with the majority of the response being directed against one or two immunodominant epitopes. This is yet another difference with HIV-1 infection, where a greater breadth of response is thought to be important as HIV-1 readily escapes the initial CD8+ T-cell response, leading to the disappearance of these T-cell populations. Selective pressure analyses comparing HIV-2 and HIV-1 capsid sequences from Caió again showed greater negative selective pressure on the HIV-2 capsid. Furthermore, immunodominant CD8+ T-cell epitopes in the HIV-2 capsid show little or no variability in Caió subjects, despite the presence of potent CD8+ T-cell IFN-γ responses targeting these epitopes [174]. This scenario, therefore, is very similar to that of HIV-2 Env and Nabs, where either functional constraints on the HIV-2 capsid and/or restricted viral replication could limit the ability of HIV-2 to escape from CD8+ T-cell responses. It is possible that the fitness costs associated with mutations within these epitopes is too great for a virus with poor replication kinetics to tolerate. Unlike in HIV-1 infection, the first HIV-2 Gag-specific CD8+ T-cell responses may therefore persist and continue to aid viral control for decades. Longitudinal studies are required to confirm whether escape from CD8+ T-cell responses is in fact a rare occurrence in HIV-2 infected subjects. Studies in the Caió cohort have also revealed that HIV-2 specific CD8+ T-cells have exceptionally high functional avidity and are able to respond to small antigen quantities [174]. Whether the lack of escape within CD8+ T-cell targeted epitopes contribute to the development of this high avidity over time is worth considering. A longitudinal study of viral sequence changes and both Nab and T-cell responses

during acute HIV-2 infection could provide clarity on a number of issues raised above. With the falling incidence of HIV-2 infection in West Africa, it is difficult to imagine such a study would be possible without dedicating considerable resources to establishing frequent HIV-2 screening programmes, in many countries, over a long period of time. Regular screening of HIV-2 discordant couples is, however, a potential option for a more targeted approach. Little is also known about HIV-2 latency in chronic infection. Do HIV-2 ECs have long lived latent viral populations which undergo little replication and evolution or is there constant viral turnover below the limit of detection of the viral load assay used so far? A longitudinal study sequencing HIV-2 from ECs would help answer this question.

Despite the previous and our current findings from Caió implying the importance of T-cell responses in HIV-2 control, it is still difficult to be certain whether the association between low VL and potent HIV-2 Gag-specific CD8+ responses is cause or consequence. This is not unique to HIV-2, but is a problem faced by all HIV-1 studies exploring correlates of protective immunity in natural infection. The lack of association with Nab responses and lower HIV-2 VL, but clear inverse correlation between both the magnitude and quality of the HIV-2 Gag-specific CD8+ response and viraemia, favours a causative role for the CD8+ T-cell response in HIV-2 viral control. Nevertheless, if HIV-2 Gag specific CD8+ responses could be manipulated or boosted in viraemic HIV-2 subjects to convert these individuals into elite controllers, this would provide further proof for the central role of T-cells in controlling HIV. The cumulative data now available from studies in Caió on HIV-2 specific T-cell responses provide a platform for designing immuno-therapeutic strategies to achieve this target. This endevour should be made a priority in future HIV-2 research, as it has the greatest potential to aid the search for an HIV-1

vaccine by providing proof of concept for vaccine-induced HIV-1 control, as well as providing an important intervention for clinicians faced with treating intractable HIV-2 infection.

REFERENCES

1. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J*, et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983,**220**:868-871.
2. UNAIDS Report on the global HIV/AIDS epidemic. Geneva. 2008.
3. UNAIDS Data Tables. Geneva. 2011.
4. Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira MO*, et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science* 1986,**233**:343-346.
5. Quinn TC. Population migration and the spread of types 1 and 2 human immunodeficiency viruses. *Proc Natl Acad Sci U S A* 1994,**91**:2407-2414.
6. Peterson K, Jallow S, Rowland-Jones SL, de Silva TI. Antiretroviral Therapy for HIV-2 Infection: Recommendations for Management in Low-Resource Settings. *AIDS Res Treat* 2011,**2011**:463704.
7. Gao F, Yue L, White AT, Pappas PG, Barchue J, Hanson AP*, et al.* Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. *Nature* 1992,**358**:495-499.
8. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* 1989,**339**:389-392.
9. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S. Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* 1990,**345**:356-359.
10. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML*, et al.* Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 2006,**313**:523-526.
11. Sharp PM, Shaw GM, Hahn BH. Simian immunodeficiency virus infection of chimpanzees. *J Virol* 2005,**79**:3891-3902.
12. Takehisa J, Kraus MH, Ayouba A, Bailes E, Van Heuverswyn F, Decker JM*, et al.* Origin and biology of simian immunodeficiency virus in wild-living western gorillas. *J Virol* 2009,**83**:1635-1648.
13. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. *Science* 2000,**287**:607-614.
14. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF*, et al.* Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* 1999,**397**:436-441.
15. Vallari A, Bodelle P, Ngansop C, Makamche F, Ndembi N, Mbanya D*, et al.* Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low. *AIDS Res Hum Retroviruses* 2010,**26**:109-115.
16. Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF*, et al.* Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. *Virology* 2007,**368**:155-171.
17. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V*, et al.* A new human immunodeficiency virus derived from gorillas. *Nat Med* 2009,**15**:871-872.
18. Peeters M, Gueye A, Mboup S, Bibollet-Ruche F, Ekaza E, Mulanga C*, et al.* Geographical distribution of HIV-1 group O viruses in Africa. *AIDS* 1997,**11**:493-498.

19. Pepin J. *The Origin of AIDS*: Cambride University Press; 2011.
20. Lemey P, Pybus OG, Rambaut A, Drummond AJ, Robertson DL, Roques P, *et al.* The molecular population genetics of HIV-1 group O. *Genetics* 2004,**167**:1059-1068.
21. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, *et al.* Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 2008,**455**:661-664.
22. Sauter D, Hue S, Petit SJ, Plantier JC, Towers GJ, Kirchhoff F, *et al.* HIV-1 Group P is unable to antagonize human tetherin by Vpu, Env or Nef. *Retrovirology* 2011,**8**:103.
23. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 1995,**373**:123-126.
24. Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science* 1988,**242**:1168-1171.
25. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, *et al.* Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 1995,**373**:117-122.
26. Chen J, Powell D, Hu WS. High frequency of genetic recombination is a common feature of primate lentivirus replication. *J Virol* 2006,**80**:9651-9658.
27. HIV Databases. HIV Sequence database. http://www.hiv.lanl.gov (accessed 19th March 2012). In.
28. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* 2011,**11**:45-56.
29. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, *et al.* HIV-1 nomenclature proposal. *Science* 2000,**288**:55-56.
30. Geretti AM. HIV-1 subtypes: epidemiology and significance for HIV management. *Curr Opin Infect Dis* 2006,**19**:1-7.
31. Taylor BS, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med* 2008,**359**:1965-1966.
32. Chen Z, Telfier P, Gettie A, Reed P, Zhang L, Ho DD, *et al.* Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *J Virol* 1996,**70**:3617-3627.
33. Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, *et al.* Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. *AIDS Res Hum Retroviruses* 2004,**20**:666-672.
34. Damond F, Descamps D, Farfara I, Telles JN, Puyeo S, Campa P, *et al.* Quantification of proviral load of human immunodeficiency virus type 2 subtypes A and B using real-time PCR. *J Clin Microbiol* 2001,**39**:4264-4268.
35. Peeters M, Toure-Kane C, Nkengasong JN. Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. *AIDS* 2003,**17**:2547-2560.
36. Pieniazek D, Ellenberger D, Janini LM, Ramos AC, Nkengasong J, Sassan-Morokro M, *et al.* Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast. *AIDS Res Hum Retroviruses* 1999,**15**:603-608.

37. Chen Z, Luckay A, Sodora DL, Telfer P, Reed P, Gettie A*, et al.* Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J Virol* 1997,**71**:3953-3960.
38. Santiago ML, Range F, Keele BF, Li Y, Bailes E, Bibollet-Ruche F*, et al.* Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *J Virol* 2005,**79**:12515-12527.
39. Smith SM, Christian D, de Lame V, Shah U, Austin L, Gautam R*, et al.* Isolation of a new HIV-2 group in the US. *Retrovirology* 2008,**5**:103.
40. Ayouba A, Akoua-Koffi C, Adjogoua E, Calvignac S, Esteban A, Locatelli S*, et al.* Identification of a New HIV-2 Lineage in Rural Cote d'Ivoire: An Additional Cross-species Transmission from SIVsmm from Sooty Mangabeys to Humans. *19th Conference on Retroviruses and Opportunistic Infections, Seattle, USA* 2012,**Abstract 62LB**.
41. Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S*, et al.* HIV-2 CRF01_AB: first circulating recombinant form of HIV-2. *J Acquir Immune Defic Syndr* 2010,**54**:241-247.
42. Gao F, Yue L, Robertson DL, Hill SC, Hui H, Biggar RJ*, et al.* Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J Virol* 1994,**68**:7433-7447.
43. Apetrei C, Metzger MJ, Richardson D, Ling B, Telfer PT, Reed P*, et al.* Detection and partial characterization of simian immunodeficiency virus SIVsm strains from bush meat samples from rural Sierra Leone. *J Virol* 2005,**79**:2631-2636.
44. Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme AM. Tracing the origin and history of the HIV-2 epidemic. *Proc Natl Acad Sci U S A* 2003,**100**:6588-6592.
45. Valadas E, Franca L, Sousa S, Antunes F. 20 years of HIV-2 infection in Portugal: trends and changes in epidemiology. *Clin Infect Dis* 2009,**48**:1166-1167.
46. Cortes E, Detels R, Aboulafia D, Li XL, Moudgil T, Alam M*, et al.* HIV-1, HIV-2, and HTLV-I infection in high-risk groups in Brazil. *N Engl J Med* 1989,**320**:953-958.
47. Kannangai R, Nair SC, Sridharan G, Prasannakumar S, Daniel D. Frequency of HIV type 2 infections among blood donor population from India: a 10-year experience. *Indian J Med Microbiol* 2010,**28**:111-113.
48. Schim van der Loeff MF, Aaby P. Towards a better understanding of the epidemiology of HIV-2. *Aids* 1999,**13 Suppl A**:S69-84.
49. Poulsen AG, Aaby P, Gottschau A, Kvinesdal BB, Dias F, Molbak K*, et al.* HIV-2 infection in Bissau, West Africa, 1987-1989: incidence, prevalences, and routes of transmission. *J Acquir Immune Defic Syndr* 1993,**6**:941-948.
50. Aaby P, Ariyoshi K, Buckner M, Jensen H, Berry N, Wilkins A*, et al.* Age of wife as a major determinant of male-to-female transmission of HIV-2 infection: a community study from rural West Africa. *Aids* 1996,**10**:1585-1590.
51. van Tienen C, Schim van der Loeff MF, Zaman SM, Vincent T, Sarge-Njie R, Peterson I*, et al.* Two distinct epidemics: the rise of HIV-1 and decline of

HIV-2 infection between 1990 and 2007 in rural Guinea-Bissau. *J Acquir Immune Defic Syndr* 2009.

52. da Silva ZJ, Oliveira I, Andersen A, Dias F, Rodrigues A, Holmgren B*, et al.* Changes in prevalence and incidence of HIV-1, HIV-2 and dual infections in urban areas of Bissau, Guinea-Bissau: is HIV-2 disappearing? *Aids* 2008,**22**:1195-1202.
53. Hamel DJ, Sankale JL, Eisen G, Meloni ST, Mullins C, Gueye-Ndiaye A*, et al.* Twenty years of prospective molecular epidemiology in Senegal: changes in HIV diversity. In: *AIDS Res Hum Retroviruses*. 2007/10/27 ed; 2007. pp. 1189-1196.
54. O'Donovan D, Ariyoshi K, Milligan P, Ota M, Yamuah L, Sarge-Njie R*, et al.* Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. MRC/Gambia Government/University College London Medical School working group on mother-child transmission of HIV. *Aids* 2000,**14**:441-448.
55. Hawes SE, Sow PS, Stern JE, Critchlow CW, Gottlieb GS, Kiviat NB. Lower levels of HIV-2 than HIV-1 in the female genital tract: correlates and longitudinal assessment of viral shedding. *AIDS* 2008,**22**:2517-2525.
56. Gottlieb GS, Hawes SE, Agne HD, Stern JE, Critchlow CW, Kiviat NB*, et al.* Lower levels of HIV RNA in semen in HIV-2 compared with HIV-1 infection: implications for differences in transmission. *AIDS* 2006,**20**:895-900.
57. Pepin J, Plamondon M, Alves AC, Beaudet M, Labbe AC. Parenteral transmission during excision and treatment of tuberculosis and trypanosomiasis may be responsible for the HIV-2 epidemic in Guinea-Bissau. *AIDS* 2006,**20**:1303-1311.
58. Poulsen AG, Aaby P, Jensen H, Dias F. Risk factors for HIV-2 seropositivity among older people in Guinea-Bissau. A search for the early history of HIV-2 infection. *Scand J Infect Dis* 2000,**32**:169-175.
59. Barin F, M'Boup S, Denis F, Kanki P, Allan JS, Lee TH*, et al.* Serological evidence for virus related to simian T-lymphotropic retrovirus III in residents of west Africa. *Lancet* 1985,**2**:1387-1389.
60. Brun-Vezinet F, Rey MA, Katlama C, Girard PM, Roulot D, Yeni P*, et al.* Lymphadenopathy-associated virus type 2 in AIDS and AIDS-related complex. Clinical and virological features in four patients. *Lancet* 1987,**1**:128-132.
61. Clavel F, Mansinho K, Chamaret S, Guetard D, Favier V, Nina J*, et al.* Human immunodeficiency virus type 2 infection associated with AIDS in West Africa. *N Engl J Med* 1987,**316**:1180-1185.
62. Martinez-Steele E, Awasana AA, Corrah T, Sabally S, van der Sande M, Jaye A*, et al.* Is HIV-2- induced AIDS different from HIV-1-associated AIDS? Data from a West African clinic. *Aids* 2007,**21**:317-324.
63. Ariyoshi K, Schim van der Loeff M, Cook P, Whitby D, Corrah T, Jaffar S*, et al.* Kaposi's sarcoma in the Gambia, West Africa is less frequent in human immunodeficiency virus type 2 than in human immunodeficiency virus type 1 infection despite a high prevalence of human herpesvirus 8. *J Hum Virol* 1998,**1**:193-199.
64. Marlink R, Kanki P, Thior I, Travers K, Eisen G, Siby T*, et al.* Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* 1994,**265**:1587-1590.
65. Marlink RG, Ricard D, M'Boup S, Kanki PJ, Romet-Lemonne JL, N'Doye I*, et al.* Clinical, hematologic, and immunologic cross-sectional evaluation of

individuals exposed to human immunodeficiency virus type-2 (HIV-2). *AIDS Res Hum Retroviruses* 1988,**4**:137-148.
66. Poulsen AG, Aaby P, Larsen O, Jensen H, Naucler A, Lisse IM*, et al.* 9-year HIV-2-associated mortality in an urban community in Bissau, west Africa. *Lancet* 1997,**349**:911-914.
67. Ricard D, Wilkins A, N'Gum PT, Hayes R, Morgan G, Da Silva AP*, et al.* The effects of HIV-2 infection in a rural area of Guinea-Bissau. *Aids* 1994,**8**:977-982.
68. Todd J, Balira R, Grosskurth H, Mayaud P, Mosha F, ka-Gina G*, et al.* HIV-associated adult mortality in a rural Tanzanian population. *AIDS* 1997,**11**:801-807.
69. van Tienen C, Schim van der Loeff M, Peterson I, Cotten M, Andersson S, Holmgren B*, et al.* HTLV-1 and HIV-2 infection are associated with increased mortality in a rural West African community. *PLoS One* 2011,**6**:e29026.
70. Schim van der Loeff MF, Jaffar S, Aveika AA, Sabally S, Corrah T, Harding E*, et al.* Mortality of HIV-1, HIV-2 and HIV-1/HIV-2 dually infected patients in a clinic-based cohort in The Gambia. *Aids* 2002,**16**:1775-1783.
71. Hansmann A, Schim van der Loeff MF, Kaye S, Awasana AA, Sarge-Njie R, O'Donovan D*, et al.* Baseline plasma viral load and CD4 cell percentage predict survival in HIV-1- and HIV-2-infected women in a community-based cohort in The Gambia. *J Acquir Immune Defic Syndr* 2005,**38**:335-341.
72. Barker E, Mackewicz CE, Reyes-Teran G, Sato A, Stranford SA, Fujimura SH*, et al.* Virological and immunological features of long-term human immunodeficiency virus-infected individuals who have remained asymptomatic compared with those who have progressed to acquired immunodeficiency syndrome. *Blood* 1998,**92**:3105-3114.
73. Ariyoshi K, Jaffar S, Alabi AS, Berry N, Schim van der Loeff M, Sabally S*, et al.* Plasma RNA viral load predicts the rate of CD4 T cell decline and death in HIV-2-infected patients in West Africa. *Aids* 2000,**14**:339-344.
74. Berry N, Ariyoshi K, Jaffar S, Sabally S, Corrah T, Tedder R*, et al.* Low peripheral blood viral HIV-2 RNA in individuals with high CD4 percentage differentiates HIV-2 from HIV-1 infection. *J Hum Virol* 1998,**1**:457-468.
75. Schim van der Loeff MF, Larke N, Kaye S, Berry N, Ariyoshi K, Alabi A*, et al.* Undetectable plasma viral load predicts normal survival in HIV-2-infected people in a West African village. *Retrovirology* 2010,**7**:46.
76. Berry N, Jaffar S, Schim van der Loeff M, Ariyoshi K, Harding E, N'Gom PT*, et al.* Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. *AIDS Res Hum Retroviruses* 2002,**18**:1167-1173.
77. Gottlieb GS, Sow PS, Hawes SE, Ndoye I, Redman M, Coll-Seck AM*, et al.* Equal plasma viral loads predict a similar rate of CD4+ T cell decline in human immunodeficiency virus (HIV) type 1- and HIV-2-infected individuals from Senegal, West Africa. *J Infect Dis* 2002,**185**:905-914.
78. Besnier JM, Barin F, Baillou A, Liard F, Choutet P, Goudeau A. Symptomatic HIV-2 primary infection. *Lancet* 1990,**335**:798.
79. Andersson S, Norrgren H, da Silva Z, Biague A, Bamba S, Kwok S*, et al.* Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. *Arch Intern Med* 2000,**160**:3286-3293.

80. Ota MO, O'Donovan D, Alabi AS, Milligan P, Yamuah LK, N'Gom PT, *et al.* Maternal HIV-1 and HIV-2 infection and child survival in The Gambia. *Aids* 2000,**14**:435-439.
81. Barroso H, Araujo F, Gomes MH, Mota-Miranda A, Taveira N. Phylogenetic demonstration of two cases of perinatal human immunodeficiency virus type 2 infection diagnosed in adulthood. *AIDS Res Hum Retroviruses* 2004,**20**:1373-1376.
82. Guyader M, Emerman M, Sonigo P, Clavel F, Montagnier L, Alizon M. Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature* 1987,**326**:662-669.
83. Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 2008,**451**:425-430.
84. Le Tortorec A, Neil SJ. Antagonism to and intracellular sequestration of human tetherin by the human immunodeficiency virus type 2 envelope glycoprotein. *J Virol* 2009,**83**:11966-11978.
85. Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 2011,**474**:654-657.
86. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, *et al.* SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol* 2012,**13**:223-228.
87. Duvall MG, Lore K, Blaak H, Ambrozak DA, Adams WC, Santos K, *et al.* Dendritic cells are less susceptible to human immunodeficiency virus type 2 (HIV-2) infection than to HIV-1 infection. *J Virol* 2007,**81**:13486-13498.
88. http://creativecommons.org/licenses/by/2.0.
89. Kong LI, Lee SW, Kappes JC, Parkin JS, Decker D, Hoxie JA, *et al.* West African HIV-2-related human retrovirus with attenuated cytopathicity. *Science* 1988,**240**:1525-1529.
90. Schramm B, Penn ML, Palacios EH, Grant RM, Kirchhoff F, Goldsmith MA. Cytopathicity of human immunodeficiency virus type 2 (HIV-2) in human lymphoid tissue is coreceptor dependent and comparable to that of HIV-1. *J Virol* 2000,**74**:9594-9600.
91. McKnight A, Dittmar MT, Moniz-Periera J, Ariyoshi K, Reeves JD, Hibbitts S, *et al.* A broad range of chemokine receptors are used by primary isolates of human immunodeficiency virus type 2 as coreceptors with CD4. *J Virol* 1998,**72**:4065-4071.
92. Blaak H, Boers PH, Gruters RA, Schuitemaker H, van der Ende ME, Osterhaus AD. CCR5, GPR15, and CXCR6 are major coreceptors of human immunodeficiency virus type 2 variants isolated from individuals with and without plasma viremia. *J Virol* 2005,**79**:1686-1700.
93. Popper SJ, Sarr AD, Travers KU, Gueye-Ndiaye A, Mboup S, Essex ME, *et al.* Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J Infect Dis* 1999,**180**:1116-1121.
94. Popper SJ, Sarr AD, Gueye-Ndiaye A, Mboup S, Essex ME, Kanki PJ. Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. *J Virol* 2000,**74**:1554-1557.
95. Alabi AS, Jaffar S, Ariyoshi K, Blanchard T, Schim van der Loeff M, Awasana AA, *et al.* Plasma viral load, CD4 cell percentage, HLA and survival of HIV-1, HIV-2, and dually infected Gambian patients. *Aids* 2003,**17**:1513-1520.

96. MacNeil A, Sarr AD, Sankale JL, Meloni ST, Mboup S, Kanki P. Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection. *J Virol* 2007,**81**:5325-5330.
97. MacNeil A, Sankale JL, Meloni ST, Sarr AD, Mboup S, Kanki P. Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. *J Virol* 2006,**80**:7316-7321.
98. Soares RS, Tendeiro R, Foxall RB, Baptista AP, Cavaleiro R, Gomes P, *et al.* Cell-associated viral burden provides evidence of ongoing viral replication in aviremic HIV-2-infected patients. *J Virol* 2011,**85**:2429-2438.
99. Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ. The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. *J Virol* 2005,**79**:8979-8990.
100. Blaak H, van der Ende ME, Boers PH, Schuitemaker H, Osterhaus AD. In vitro replication capacity of HIV-2 variants from long-term aviremic individuals. *Virology* 2006,**353**:144-154.
101. MacNeil A, Sankale JL, Meloni ST, Sarr AD, Mboup S, Kanki P. Long-term intrapatient viral evolution during HIV-2 infection. *J Infect Dis* 2007,**195**:726-733.
102. Skar H, Borrego P, Wallstrom TC, Mild M, Marcelino JM, Barroso H, *et al.* HIV-2 genetic evolution in patients with advanced disease is faster than that in matched HIV-1 patients. *J Virol* 2010,**84**:7412-7415.
103. Grassly NC, Xiang Z, Ariyoshi K, Aaby P, Jensen H, Schim van der Loeff M, *et al.* Mortality among human immunodeficiency virus type 2-positive villagers in rural Guinea-Bissau is correlated with viral genotype. *J Virol* 1998,**72**:7895-7899.
104. Onyango CO, Leligdowicz A, Yokoyama M, Sato H, Song H, Nakayama EE, *et al.* HIV-2 capsids distinguish high and low virus load patients in a West African community cohort. *Vaccine* 2010,**28 Suppl 2**:B60-67.
105. Song H, Nakayama EE, Yokoyama M, Sato H, Levy JA, Shioda T. A single amino acid of the human immunodeficiency virus type 2 capsid affects its replication in the presence of cynomolgus monkey and human TRIM5alphas. *J Virol* 2007,**81**:7280-7285.
106. Giorgi JV, Hultin LE, McKeating JA, Johnson TD, Owens B, Jacobson LP, *et al.* Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *J Infect Dis* 1999,**179**:859-870.
107. Jaffar S, Van der Loeff MS, Eugen-Olsen J, Vincent T, Sarje-Njie R, Ngom P, *et al.* Immunological predictors of survival in HIV type 2-infected rural villagers in Guinea-Bissau. *AIDS Res Hum Retroviruses* 2005,**21**:560-564.
108. Sousa AE, Carneiro J, Meier-Schellersheim M, Grossman Z, Victorino RM. CD4 T cell depletion is linked directly to immune activation in the pathogenesis of HIV-1 and HIV-2 but only indirectly to the viral load. *J Immunol* 2002,**169**:3400-3406.
109. Leligdowicz A, Feldmann J, Jaye A, Cotten M, Dong T, McMichael A, *et al.* Direct relationship between virus load and systemic immune activation in HIV-2 infection. *J Infect Dis* 2010,**201**:114-122.
110. Schindler M, Munch J, Kutsch O, Li H, Santiago ML, Bibollet-Ruche F, *et al.* Nef-mediated suppression of T cell activation was lost in a lentiviral lineage that gave rise to HIV-1. *Cell* 2006,**125**:1055-1067.

111. Feldmann J, Leligdowicz A, Jaye A, Dong T, Whittle H, Rowland-Jones SL. T-cell receptor (TCR) down-regulation by HIV-2 nef does not protect against disease progression. *J Virol* 2009.
112. Khalid M, Yu H, Sauter D, Usmani SM, Schmokel J, Feldman J*, et al.* Efficient Nef-mediated downmodulation of TCR-CD3 and CD28 is associated with high CD4+ T cell counts in viremic HIV-2 infection. *J Virol* 2012,**86**:4906-4920.
113. Michel P, Balde AT, Roussilhon C, Aribot G, Sarthou JL, Gougeon ML. Reduced immune activation and T cell apoptosis in human immunodeficiency virus type 2 compared with type 1: correlation of T cell apoptosis with beta2 microglobulin concentration and disease evolution. *J Infect Dis* 2000,**181**:64-75.
114. Jaleco AC, Covas MJ, Victorino RM. Analysis of lymphocyte cell death and apoptosis in HIV-2-infected patients. *Clin Exp Immunol* 1994,**98**:185-189.
115. Machuca A, Ding L, Taffs R, Lee S, Wood O, Hu J*, et al.* HIV type 2 primary isolates induce a lower degree of apoptosis "in vitro" compared with HIV type 1 primary isolates. *AIDS Res Hum Retroviruses* 2004,**20**:507-512.
116. Gautier D, Beq S, Cortesao CS, Sousa AE, Cheynier R. Efficient thymopoiesis contributes to the maintenance of peripheral CD4 T cells during chronic human immunodeficiency virus type 2 infection. *J Virol* 2007,**81**:12685-12688.
117. Albuquerque AS, Cortesao CS, Foxall RB, Soares RS, Victorino RM, Sousa AE. Rate of increase in circulating IL-7 and loss of IL-7Ralpha expression differ in HIV-1 and HIV-2 infections: two lymphopenic diseases with similar hyperimmune activation but distinct outcomes. *J Immunol* 2007,**178**:3252-3259.
118. Mattapallil JJ, Douek DC, Hill B, Nishimura Y, Martin M, Roederer M. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* 2005,**434**:1093-1097.
119. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, Rao S*, et al.* Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med* 2006.
120. Nowroozalizadeh S, Mansson F, da Silva Z, Repits J, Dabo B, Pereira C*, et al.* Microbial translocation correlates with the severity of both HIV-1 and HIV-2 infections. *J Infect Dis* 2010,**201**:1150-1154.
121. Pandrea IV, Gautam R, Ribeiro RM, Brenchley JM, Butler IF, Pattison M*, et al.* Acute loss of intestinal CD4+ T cells is not predictive of simian immunodeficiency virus virulence. *J Immunol* 2007,**179**:3035-3046.
122. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD*, et al.* The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* 2010,**330**:1551-1557.
123. Diouf K, Sarr AD, Eisen G, Popper S, Mboup S, Kanki P. Associations between MHC class I and susceptibility to HIV-2 disease progression. *J Hum Virol* 2002,**5**:1-7.
124. Yindom LM, Leligdowicz A, Martin MP, Gao X, Qi Y, Zaman SM*, et al.* Influence of HLA class I and HLA-KIR compound genotypes on HIV-2 infection and markers of disease progression in a Manjako community in West Africa. *J Virol* 2010,**84**:8202-8208.
125. Burton DR. Antibodies, viruses and vaccines. *Nat Rev Immunol* 2002,**2**:706-713.

126. Moog C, Fleury HJ, Pellegrin I, Kirn A, Aubertin AM. Autologous and heterologous neutralizing antibody responses following initial seroconversion in human immunodeficiency virus type 1-infected individuals. *J Virol* 1997,**71**:3734-3741.
127. Johnson WE, Desrosiers RC. Viral persistance: HIV's strategies of immune system evasion. *Annu Rev Med* 2002,**53**:499-518.
128. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X*, et al.* Antibody neutralization and escape by HIV-1. *Nature* 2003,**422**:307-312.
129. Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* 2003,**100**:4144-4149.
130. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 1998,**393**:648-659.
131. Wyatt R, Kwong PD, Desjardins E, Sweet RW, Robinson J, Hendrickson WA*, et al.* The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 1998,**393**:705-711.
132. Parren PW, Burton DR, Sattentau QJ. HIV-1 antibody--debris or virion? *Nat Med* 1997,**3**:366-367.
133. Stamatatos L, Morris L, Burton DR, Mascola JR. Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine? *Nat Med* 2009,**15**:866-870.
134. Alter G, Moody MA. The humoral response to HIV-1: new insights, renewed focus. *J Infect Dis* 2010,**202 Suppl 2**:S315-322.
135. Mahalanabis M, Jayaraman P, Miura T, Pereyra F, Chester EM, Richardson B*, et al.* Continuous viral escape and selection by autologous neutralizing antibodies in drug-naive human immunodeficiency virus controllers. *J Virol* 2009,**83**:662-672.
136. Piantadosi A, Panteleeff D, Blish CA, Baeten JM, Jaoko W, McClelland RS*, et al.* Breadth of neutralizing antibody response to human immunodeficiency virus type 1 is affected by factors early in infection but does not influence disease progression. *J Virol* 2009,**83**:10269-10274.
137. Gray ES, Moore PL, Choge IA, Decker JM, Bibollet-Ruche F, Li H*, et al.* Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection. *J Virol* 2007,**81**:6187-6196.
138. Bailey JR, Lassen KG, Yang HC, Quinn TC, Ray SC, Blankson JN*, et al.* Neutralizing antibodies do not mediate suppression of human immunodeficiency virus type 1 in elite suppressors or selection of plasma virus variants in patients on highly active antiretroviral therapy. *J Virol* 2006,**80**:4758-4770.
139. Weiss RA, Clapham PR, Weber JN, Whitby D, Tedder RS, O'Connor T*, et al.* HIV-2 antisera cross-neutralize HIV-1. *Aids* 1988,**2**:95-100.
140. Schulz TF, Whitby D, Hoad JG, Corrah T, Whittle H, Weiss RA. Biological and molecular variability of human immunodeficiency virus type 2 isolates from The Gambia. *J Virol* 1990,**64**:5177-5182.
141. Bjorling E, Scarlatti G, von Gegerfelt A, Albert J, Biberfeld G, Chiodi F*, et al.* Autologous neutralizing antibodies prevail in HIV-2 but not in HIV-1 infection. *Virology* 1993,**193**:528-530.
142. Shi Y, Brandin E, Vincic E, Jansson M, Blaxhult A, Gyllensten K*, et al.* Evolution of human immunodeficiency virus type 2 coreceptor usage,

autologous neutralization, envelope sequence and glycosylation. *J Gen Virol* 2005,**86**:3385-3396.
143. Thomas ER, Shotton C, Weiss RA, Clapham PR, McKnight A. CD4-dependent and CD4-independent HIV-2: consequences for neutralization. *Aids* 2003,**17**:291-300.
144. Decker JM, Bibollet-Ruche F, Wei X, Wang S, Levy DN, Wang W*, et al.* Antigenic conservation and immunogenicity of the HIV coreceptor binding site. *J Exp Med* 2005,**201**:1407-1419.
145. Rodriguez SK, Sarr AD, Macneil A, Thakore-Meloni S, Gueye-Ndiaye A, Traore I*, et al.* Comparison of heterologous neutralizing antibody responses between HIV-1 and HIV-2 infected Senegalese patients: Distinct patterns of breadth and magnitude distinguish HIV-1 and HIV-2. *J Virol* 2007.
146. Tendeiro R, Fernandes S, Foxall RB, Marcelino JM, Taveira N, Soares RS*, et al.* Memory B-cell depletion is a feature of HIV-2 infection even in the absence of detectable viremia. *AIDS* 2012, **26**:1607-1617.
147. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *J Virol* 1994,**68**:6103-6110.
148. Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, Borkowsky W*, et al.* Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol* 1994,**68**:4650-4655.
149. Ogg GS, Jin X, Bonhoeffer S, Dunbar PR, Nowak MA, Monard S*, et al.* Quantitation of HIV-1-specific cytotoxic T lymphocytes and plasma load of viral RNA. *Science* 1998,**279**:2103-2106.
150. Ogg GS, Kostense S, Klein MR, Jurriaans S, Hamann D, McMichael AJ*, et al.* Longitudinal phenotypic analysis of human immunodeficiency virus type 1-specific cytotoxic T lymphocytes: correlation with disease progression. *J Virol* 1999,**73**:9153-9160.
151. Walker BD, Chakrabarti S, Moss B, Paradis TJ, Flynn T, Durno AG*, et al.* HIV-specific cytotoxic T lymphocytes in seropositive individuals. *Nature* 1987,**328**:345-348.
152. Hansen SG, Ford JC, Lewis MS, Ventura AB, Hughes CM, Coyne-Johnson L*, et al.* Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* 2011,**473**:523-527.
153. Schmitz JE, Kuroda MJ, Santra S, Sasseville VG, Simon MA, Lifton MA*, et al.* Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* 1999,**283**:857-860.
154. Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, Vogel M*, et al.* Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. *Aids* 2003,**17**:2581-2591.
155. Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, Martino L*, et al.* HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proc Natl Acad Sci U S A* 2000,**97**:2709-2714.
156. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA*, et al.* Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat Med* 1997,**3**:212-217.
157. McMichael A, Klenerman P. HIV/AIDS. HLA leaves its footprints on HIV. *Science* 2002,**296**:1410-1411.

158. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* 2002,**296**:1439-1443.
159. Saez-Cirion A, Lacabaratz C, Lambotte O, Versmisse P, Urrutia A, Boufassa F, *et al.* HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. *Proc Natl Acad Sci U S A* 2007,**104**:6776-6781.
160. Day CL, Kiepiela P, Leslie AJ, van der Stok M, Nair K, Ismail N, *et al.* Proliferative capacity of epitope-specific CD8 T-cell responses is inversely related to viral load in chronic human immunodeficiency virus type 1 infection. *J Virol* 2007,**81**:434-438.
161. Betts MR, Nason MC, West SM, De Rosa SC, Migueles SA, Abraham J, *et al.* HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood* 2006,**107**:4781-4789.
162. Critchfield JW, Lemongello D, Walker DH, Garcia JC, Asmuth DM, Pollard RB, *et al.* Multifunctional human immunodeficiency virus (HIV) gag-specific CD8+ T-cell responses in rectal mucosa and peripheral blood mononuclear cells during chronic HIV type 1 infection. *J Virol* 2007,**81**:5460-5471.
163. Migueles SA, Laborico AC, Shupert WL, Sabbaghian MS, Rabin R, Hallahan CW, *et al.* HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nat Immunol* 2002,**3**:1061-1068.
164. Migueles SA, Osborne CM, Royce C, Compton AA, Joshi RP, Weeks KA, *et al.* Lytic granule loading of CD8+ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity* 2008,**29**:1009-1021.
165. Hersperger AR, Pereyra F, Nason M, Demers K, Sheth P, Shin LY, *et al.* Perforin expression directly ex vivo by HIV-specific CD8 T-cells is a correlate of HIV elite control. *PLoS Pathog* 2010,**6**:e1000917.
166. Gillespie GM, Pinheiro S, Sayeid-Al-Jamee M, Alabi A, Kaye S, Sabally S, *et al.* CD8+ T cell responses to human immunodeficiency viruses type 2 (HIV-2) and type 1 (HIV-1) gag proteins are distinguishable by magnitude and breadth but not cellular phenotype. *Eur J Immunol* 2005,**35**:1445-1453.
167. Jaye A, Sarge-Njie R, Schim van der Loeff M, Todd J, Alabi A, Sabally S, *et al.* No differences in cellular immune responses between asymptomatic HIV type 1- and type 2-infected Gambian patients. *J Infect Dis* 2004,**189**:498-505.
168. Zheng NN, Kiviat NB, Sow PS, Hawes SE, Wilson A, Diallo-Agne H, *et al.* Comparison of human immunodeficiency virus (HIV)-specific T-cell responses in HIV-1- and HIV-2-infected individuals in Senegal. *J Virol* 2004,**78**:13934-13942.
169. Duvall MG, Jaye A, Dong T, Brenchley JM, Alabi AS, Jeffries DJ, *et al.* Maintenance of HIV-specific CD4+ T cell help distinguishes HIV-2 from HIV-1 infection. *J Immunol* 2006,**176**:6973-6981.
170. Duvall MGPM, Ambrozak DA, Jaye A, McMichael AJ, Whittle HC, Roederer M, Rowland-Jones SL, Koup RA. Polyfunctional T cell responses are a hallmark of HIV-2 infection. *Eur J Immunol* 2008,**38**:350-363.
171. Ariyoshi K, Cham F, Berry N, Jaffar S, Sabally S, Corrah T, *et al.* HIV-2-specific cytotoxic T-lymphocyte activity is inversely related to proviral load. *Aids* 1995,**9**:555-559.
172. Leligdowicz A, Yindom LM, Onyango C, Sarge-Njie R, Alabi A, Cotten M, *et al.* Robust Gag-specific T cell responses characterize viremia control in HIV-2 infection. *J Clin Invest* 2007,**117**:3067-3074.

173. Papagno L, Spina CA, Marchant A, Salio M, Rufer N, Little S*, et al.* Immune activation and CD8+ T-cell differentiation towards senescence in HIV-1 infection. *PLoS Biol* 2004,**2**:E20.
174. Leligdowicz A, Onyango C, Yindom LM, Peng Y, Cotten M, Jaye A*, et al.* Highly avid, oligoclonal, early-differentiated antigen-specific CD8(+) T-cells in chronic HIV-2 infection. *Eur J Immunol* 2010.
175. Almeida JR, Price DA, Papagno L, Arkoub ZA, Sauce D, Bornstein E*, et al.* Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *J Exp Med* 2007,**204**:2473-2485.
176. Almeida JR, Sauce D, Price DA, Papagno L, Shin SY, Moris A*, et al.* Antigen sensitivity is a major determinant of CD8+ T-cell polyfunctionality and HIV-suppressive activity. *Blood* 2009,**113**:6351-6360.
177. Lichterfeld M, Yu XG, Mui SK, Williams KL, Trocha A, Brockman MA*, et al.* Selective depletion of high-avidity human immunodeficiency virus type 1 (HIV-1)-specific CD8+ T cells after early HIV-1 infection. *J Virol* 2007,**81**:4199-4214.
178. Witvrouw M, Pannecouque C, Van Laethem K, Desmyter J, De Clercq E, Vandamme AM. Activity of non-nucleoside reverse transcriptase inhibitors against HIV-2 and SIV. *Aids* 1999,**13**:1477-1483.
179. Isaka Y, Miki S, Kawauchi S, Suyama A, Sugimoto H, Adachi A*, et al.* A single amino acid change at Leu-188 in the reverse transcriptase of HIV-2 and SIV renders them sensitive to non-nucleoside reverse transcriptase inhibitors. *Arch Virol* 2001,**146**:743-755.
180. Witvrouw M, Pannecouque C, Switzer WM, Folks TM, De Clercq E, Heneine W. Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. *Antivir Ther* 2004,**9**:57-65.
181. Damond F, Lariven S, Roquebert B, Males S, Peytavin G, Morau G*, et al.* Virological and immunological response to HAART regimen containing integrase inhibitors in HIV-2-infected patients. *Aids* 2008,**22**:665-666.
182. Reid P, MacInnes H, Cong ME, Heneine W, Garcia-Lerma JG. Natural resistance of human immunodeficiency virus type 2 to zidovudine. *Virology* 2005,**336**:251-264.
183. Boyer PL, Sarafianos SG, Clark PK, Arnold E, Hughes SH. Why do HIV-1 and HIV-2 use different pathways to develop AZT resistance? *PLoS Pathog* 2006,**2**:e10.
184. Colson P, Henry M, Tivoli N, Gallais H, Gastaut JA, Moreau J*, et al.* Polymorphism and drug-selected mutations in the reverse transcriptase gene of HIV-2 from patients living in southeastern France. *J Med Virol* 2005,**75**:381-390.
185. Desbois D, Roquebert B, Peytavin G, Damond F, Collin G, Benard A*, et al.* In vitro phenotypic susceptibility of human immunodeficiency virus type 2 clinical isolates to protease inhibitors. *Antimicrob Agents Chemother* 2008,**52**:1545-1548.
186. Masse S, Lu X, Dekhtyar T, Lu L, Koev G, Gao F*, et al.* In vitro selection and characterization of human immunodeficiency virus type 2 with decreased susceptibility to lopinavir. *Antimicrob Agents Chemother* 2007,**51**:3075-3080.
187. Ntemgwa M, Brenner BG, Oliveira M, Moisi D, Wainberg MA. Natural polymorphisms in the human immunodeficiency virus type 2 protease can

accelerate time to development of resistance to protease inhibitors. *Antimicrob Agents Chemother* 2007,**51**:604-610.
188. Drylewicz J, Matheron S, Lazaro E, Damond F, Bonnet F, Simon S*, et al.* Comparison of viro-immunological marker changes between HIV-1 and HIV-2 infected patients in France. *AIDS* 2008,**22**:457-468.
189. Peterson I, Togun O, de Silva T, Oko F, Rowland-Jones S, Jaye A*, et al.* Mortality and immunovirological outcomes on antiretroviral therapy in HIV-1 and HIV-2-infected individuals in the Gambia. *AIDS* 2011,**25**:2167-2175.
190. Buckner M. Village women as town prostitutes. In: *Vivre et penser le sida en Afrique*. Edited by Becker C, Dozon J-P, Obbo C, Touré M. Paris: Codesria, IRD, Karthala; 1999. pp. 389-404.
191. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF*, et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol* 2008,**82**:3952-3970.
192. http://www.macrogen.com. In.
193. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. . *Nucleic Acids Symposium Series* 1999,**41**:95 - 98.
194. Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 2000,**28**:1102 - 1104.
195. Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM*, et al.* Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* 2002,**46**:1896-1905.
196. http://www.hiv.lanl.gov/content/nab-reference-strains/html/home.htm. In.
197. Reed LJM, H. A simple method of estimating fifty percent endpoints. *American Journal of Hygiene* 1938,**27**:493 - 497.
198. Draenert R, Altfeld M, Brander C, Basgoz N, Corcoran C, Wurcel AG*, et al.* Comparison of overlapping peptide sets for detection of antiviral CD8 and CD4 T cell responses. *J Immunol Methods* 2003,**275**:19-29.
199. Zloza A, Schenkel JM, Tenorio AR, Martinson JA, Jeziorczak PM, Al-Harthi L. Potent HIV-specific responses are enriched in a unique subset of CD8+ T cells that coexpresses CD4 on its surface. *Blood* 2009,**114**:3841-3853.
200. Roederer M, Nozzi JL, Nason MC. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytometry A* 2011,**79**:167-174.
201. Chattopadhyay PK, Melenhorst JJ, Ladell K, Gostick E, Scheinberg P, Barrett AJ*, et al.* Techniques to improve the direct ex vivo detection of low frequency antigen-specific CD8+ T cells with peptide-major histocompatibility complex class I tetramers. *Cytometry A* 2008,**73**:1001-1009.
202. GraphPad Software, San Diego California, USA, www.graphpad.com.
203. Gao F, Robertson DL, Carruthers CD, Morrison SG, Jian B, Chen Y*, et al.* A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1. *J Virol* 1998,**72**:5680-5698.

204. Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC. Human immunodeficiency virus. Phylogeny and the origin of HIV-1. *Nature* 2001,**410**:1047-1048.
205. Ndembi N, Takehisa J, Zekeng L, Kobayashi E, Ngansop C, Songok EM*, et al.* Genetic diversity of HIV type 1 in rural eastern Cameroon. *J Acquir Immune Defic Syndr* 2004,**37**:1641-1650.
206. Tebit DM, Zekeng L, Kaptue L, Salminen M, Krausslich HG, Herchenroder O. Genotypic and phenotypic analysis of HIV type 1 primary isolates from western Cameroon. *AIDS Res Hum Retroviruses* 2002,**18**:39-48.
207. Vidal N, Peeters M, Mulanga-Kabeya C, Nzilambi N, Robertson D, Ilunga W*, et al.* Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J Virol* 2000,**74**:10498-10507.
208. Zhu T, Korber BT, Nahmias AJ, Hooper E, Sharp PM, Ho DD. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature* 1998,**391**:594-597.
209. National AIDS Secretariat of The Gambia.
210. van der Loeff MF, Awasana AA, Sarge-Njie R, van der Sande M, Jaye A, Sabally S*, et al.* Sixteen years of HIV surveillance in a West African research clinic reveals divergent epidemic trends of HIV-1 and HIV-2. *Int J Epidemiol* 2006,**35**:1322-1328.
211. Esbjornsson J, Mild M, Mansson F, Norrgren H, Medstrand P. HIV-1 molecular epidemiology in Guinea-Bissau, West Africa: origin, demography and migrations. *PLoS One* 2011,**6**:e17025.
212. Andersson S, Norrgren H, Dias F, Biberfeld G, Albert J. Molecular characterization of human immunodeficiency virus (HIV)-1 and -2 in individuals from guinea-bissau with single or dual infections: predominance of a distinct HIV-1 subtype A/G recombinant in West Africa. *Virology* 1999,**262**:312-320.
213. Kaleebu P, French N, Mahe C, Yirrell D, Watera C, Lyagoba F*, et al.* Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1-positive persons in Uganda. *J Infect Dis* 2002,**185**:1244-1250.
214. Invernizzi CF, Coutsinos D, Oliveira M, Moisi D, Brenner BG, Wainberg MA. Signature nucleotide polymorphisms at positions 64 and 65 in reverse transcriptase favor the selection of the K65R resistance mutation in HIV-1 subtype C. *J Infect Dis* 2009,**200**:1202-1206.
215. Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008). *J Int AIDS Soc* 2009,**12**:11.
216. Barouch DH, O'Brien KL, Simmons NL, King SL, Abbink P, Maxfield LF*, et al.* Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nat Med* 2010,**16**:319-323.
217. Santra S, Liao HX, Zhang R, Muldoon M, Watson S, Fischer W*, et al.* Mosaic vaccines elicit CD8+ T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys. *Nat Med* 2010,**16**:324-328.
218. Meloni ST, Kim B, Sankale JL, Hamel DJ, Tovanabutra S, Mboup S*, et al.* Distinct human immunodeficiency virus type 1 subtype A virus circulating in West Africa: sub-subtype A3. *J Virol* 2004,**78**:12438-12445.

219. Meloni ST, Sankale JL, Hamel DJ, Eisen G, Gueye-Ndiaye A, Mboup S*, et al.* Molecular epidemiology of human immunodeficiency virus type 1 sub-subtype A3 in Senegal from 1988 to 2001. *J Virol* 2004,**78**:12455-12461.
220. ClustalW2 E. http://www.ebi.ac.uk/Tools/clustalw2/index.html. In.
221. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H*, et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* 2007,**23**:2947-2948.
222. Se-Al. http://tree.bio.ed.ac.uk/software/seal/.
223. Swofford DL. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts. In; 2003.
224. Felsenstein J. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet* 1973,**25**:471-492.
225. Yang Z. Estimating the pattern of nucleotide substitution. *J Mol Evol* 1994,**39**:105-111.
226. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003,**19**:1572-1574.
227. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S*, et al.* MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Syst Biol* 2012.
228. BEAST. http://beast.bio.ed.ac.uk.
229. Tracer. http://tree.bio.ed.ac.uk/software/tracer/.
230. Figtree. http://tree.bio.ed.ac.uk/software/figtree.
231. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG*, et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999,**73**:152-160.
232. Abecasis AB, Wang Y, Libin P, Imbrechts S, de Oliveira T, Camacho RJ*, et al.* Comparative performance of the REGA subtyping tool version 2 versus version 1. *Infect Genet Evol* 2010,**10**:380-385.
233. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res* 2003,**31**:298-303.
234. Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. *Nucleic Acids Res* 2004,**32**:W654-659.
235. NCBI-BLAST. http://blast.ncbi.nlm.nih.gov/Blast.cgi.
236. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009,**26**:1641-1650.
237. The Gambia Atlas of Population and Housing Census. The Gambia Bureau of Statistics (GBoS). In; 2003.
238. Carr JK, Salminen MO, Albert J, Sanders-Buell E, Gotte D, Birx DL*, et al.* Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants. *Virology* 1998,**247**:22-31.
239. Bobkov A, Cheingsong-Popov R, Salminen M, McCutchan F, Louwagie J, Ariyoshi K*, et al.* Complex mosaic structure of the partial envelope sequence from a Gambian HIV type 1 isolate. *AIDS Res Hum Retroviruses* 1996,**12**:169-171.
240. Cham F, Heyndrickx L, Janssens W, Van der Auwera G, Vereecken K, De Houwer K*, et al.* Study of HIV type 1 gag/env variability in The Gambia,

using a multiplex DNA polymerase chain reaction. *AIDS Res Hum Retroviruses* 2000,**16**:1915-1919.
241. Binley JM, Wrin T, Korber B, Zwick MB, Wang M, Chappey C*, et al.* Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 2004,**78**:13232-13252.
242. Novitsky VA, Gaolekwe S, McLane MF, Ndung'u TP, Foley BT, Vannberg F*, et al.* HIV type 1 A/J recombinant with a pronounced pol gene mosaicism. *AIDS Res Hum Retroviruses* 2000,**16**:1015-1020.
243. Njai HF, Gali Y, Vanham G, Clybergh C, Jennes W, Vidal N*, et al.* The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. *Retrovirology* 2006,**3**:40.
244. Abraha A, Nankya IL, Gibson R, Demers K, Tebit DM, Johnston E*, et al.* CCR5- and CXCR4-tropic subtype C human immunodeficiency virus type 1 isolates have a lower level of pathogenic fitness than other dominant group M subtypes: implications for the epidemic. *J Virol* 2009,**83**:5592-5605.
245. Tscherning C, Alaeus A, Fredriksson R, Bjorndal A, Deng H, Littman DR*, et al.* Differences in chemokine coreceptor usage between genetic subtypes of HIV-1. *Virology* 1998,**241**:181-188.
246. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M*, et al.* Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: underrepresentation of X4 variants. *J Virol* 1999,**73**:6271-6281.
247. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F*, et al.* Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J Infect Dis* 2008,**197**:707-713.
248. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L*, et al.* HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J Infect Dis* 2007,**195**:1177-1180.
249. Kanki PJ, Hamel DJ, Sankale JL, Hsieh C, Thior I, Barin F*, et al.* Human immunodeficiency virus type 1 subtypes differ in disease progression. *J Infect Dis* 1999,**179**:68-73.
250. Laurent C, Bourgeois A, Faye MA, Mougnutou R, Seydi M, Gueye M*, et al.* No difference in clinical progression between patients infected with the predominant human immunodeficiency virus type 1 circulating recombinant form (CRF) 02_AG strain and patients not infected with CRF02_AG, in Western and West-Central Africa: a four-year prospective multicenter study. *J Infect Dis* 2002,**186**:486-492.
251. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E*, et al.* Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. *Clin Infect Dis* 2009,**48**:1296-1305.
252. Scherrer AU, Ledergerber B, von Wyl V, Boni J, Yerly S, Klimkait T*, et al.* Improved virological outcome in White patients infected with HIV-1 non-B subtypes compared to subtype B. *Clin Infect Dis* 2011,**53**:1143-1152.
253. Jung M, Leye N, Vidal N, Fargette D, Diop H, Toure Kane C*, et al.* The origin and evolutionary history of HIV-1 subtype C in Senegal. *PLoS One* 2012,**7**:e33579.

254. Faria NR, Suchard MA, Abecasis A, Sousa JD, Ndembi N, Bonfim I*, et al.* Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. *Infect Genet Evol* 2012,**12**:453-460.
255. Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R*, et al.* Recombination confounds the early evolutionary history of human immunodeficiency virus type 1: subtype G is a circulating recombinant form. *J Virol* 2007,**81**:8543-8551.
256. Zhang M, Foley B, Schultz AK, Macke JP, Bulla I, Stanke M*, et al.* The role of recombination in the emergence of a complex and dynamic HIV epidemic. *Retrovirology* 2010,**7**:25.
257. Ndiaye HD, Toure-Kane C, Vidal N, Niama FR, Niang-Diallo PA, Dieye T*, et al.* Surprisingly high prevalence of subtype C and specific HIV-1 subtype/CRF distribution in men having sex with men in Senegal. *J Acquir Immune Defic Syndr* 2009,**52**:249-252.
258. Pickering H, Todd J, Dunn D, Pepin J, Wilkins A. Prostitutes and their clients: a Gambian survey. *Soc Sci Med* 1992,**34**:75-88.
259. Lowndes C, Alary M, Belleau M, Kofi Bosu W, Kintin DF, Nnorom JA*, et al.* West Africa HIV/AIDS Epidemiology and Response Synthesis. *Global AIDS Monitoring and Evaluation Team (GAMET)* 2008.
260. Ariyoshi K, Berry N, Wilkins A, Ricard D, Aaby P, Naucler A*, et al.* A community-based study of human immunodeficiency virus type 2 provirus load in rural village in West Africa. *J Infect Dis* 1996,**173**:245-248.
261. Poulsen AG, Kvinesdal B, Aaby P, Molbak K, Frederiksen K, Dias F*, et al.* Prevalence of and mortality from human immunodeficiency virus type 2 in Bissau, West Africa. *Lancet* 1989,**1**:827-831.
262. Jensen ML, Dave S, Schim van der Loeff M, da Costa C, Vincent T, Leligdowicz A*, et al.* Vaccinia scars associated with improved survival among adults in rural Guinea-Bissau. *PLoS One* 2006,**1**:e101.
263. Hughes GJ, Fearnhill E, Dunn D, Lycett SJ, Rambaut A, Leigh Brown AJ. Molecular phylodynamics of the heterosexual HIV epidemic in the United Kingdom. *PLoS Pathog* 2009,**5**:e1000590.
264. Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT. Transmission Network Parameters Estimated From HIV Sequences for a Nationwide Epidemic. *J Infect Dis* 2011.
265. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ. Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLoS Med* 2008,**5**:e50.
266. Schim van der Loeff MF, Aaby P, Aryioshi K, Vincent T, Awasana AA, Da Costa C*, et al.* HIV-2 does not protect against HIV-1 infection in a rural community in Guinea-Bissau. *Aids* 2001,**15**:2303-2310.
267. Wilkins A, Ricard D, Todd J, Whittle H, Dias F, Paulo Da Silva A. The epidemiology of HIV infection in a rural area of Guinea-Bissau. *Aids* 1993,**7**:1119-1122.
268. de Silva TI, Aasa-Chapman M, Cotten M, Hue S, Robinson J, Bibollet-Ruche F*, et al.* Potent autologous and heterologous neutralizing antibody responses occur in HIV-2 infection across a broad range of infection outcomes. *J Virol* 2012,**86**:930-946.
269. Xiang Z, Ariyoshi K, Wilkins A, Dias F, Whittle H, Breuer J. HIV type 2 pathogenicity is not related to subtype in rural Guinea Bissau. *AIDS Res Hum Retroviruses* 1997,**13**:501-505.

270. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007,**7**:214.
271. Suchard MA, Weiss RE, Sinsheimer JS. Models for estimating bayes factors with applications to phylogeny and tests of monophyly. *Biometrics* 2005,**61**:665-673.
272. Lemey P, Rambaut A, Pybus OG. HIV evolutionary dynamics within and among hosts. *AIDS Rev* 2006,**8**:125-140.
273. Barroso H, Borrego P, Bartolo I, Marcelino JM, Familia C, Quintas A*, et al.* Evolutionary and structural features of the C2, V3 and C3 envelope regions underlying the differences in HIV-1 and HIV-2 biology and infection. *PLoS One* 2011,**6**:e14548.
274. Anderson RM, May RM. The population biology of the interaction between HIV-1 and HIV-2: coexistence or competitive exclusion? *AIDS* 1996,**10**:1663-1673.
275. Schmidt WP, Van Der Loeff MS, Aaby P, Whittle H, Bakker R, Buckner M*, et al.* Behaviour change and competitive exclusion can explain the diverging HIV-1 and HIV-2 prevalence trends in Guinea-Bissau. *Epidemiol Infect* 2008,**136**:551-561.
276. de Silva T, Van Tienen C, Rowland-Jones S, Cotten M. Dual infection with HIV-1 and HIV-2: double trouble or destructive interference? *HIV Therapy* 2010,**4**:305 - 323.
277. Bruhn CA, Gilbert MT. HIV-2 down, HIV-1 to go? Understanding the possibilities of treatment as prevention. *Lancet Infect Dis* 2011,**11**:260-261.
278. Montaner JS, Lima VD, Barrios R, Yip B, Wood E, Kerr T*, et al.* Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. *Lancet* 2010,**376**:532-539.
279. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N*, et al.* Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 2011,**365**:493-505.
280. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 2009,**373**:48-57.
281. van der Loeff MF, Larke N, Kaye S, Berry N, Ariyoshi K, Alabi A*, et al.* Undetectable plasma viral load predicts normal survival in HIV-2-infected people in a West African village. *Retrovirology* 2010,**7**:46.
282. Geskus R, Poulin N, Whittle H, Schim van der Loeff M. A Comparison of Progression of HIV-1 and HIV-2 Infection in The Gambia, Africa, Abstract 13_292. *13th International Workshop on HIV Observational Databababases, Lissabon, Portugal* 2009.
283. Pao D, Fisher M, Hue S, Dean G, Murphy G, Cane PA*, et al.* Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. *AIDS* 2005,**19**:85-90.
284. Faria NR, Hodges-Mameletzis I, Silva JC, Rodes B, Erasmus S, Paolucci S*, et al.* Phylogeographical footprint of colonial history in the global dispersal of human immunodeficiency virus type 2 group A. *J Gen Virol* 2012,**93**:889-899.
285. Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C*, et al.* High rates of forward transmission events after acute/early HIV-1 infection. *J Infect Dis* 2007,**195**:951-959.

286. Fisher M, Pao D, Brown AE, Sudarshi D, Gill ON, Cane P*, et al.* Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. *AIDS* 2010,**24**:1739-1747.
287. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O*, et al.* Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 2005,**191**:1403-1409.
288. Pilcher CD, Tien HC, Eron JJ, Jr., Vernazza PL, Leu SY, Stewart PW*, et al.* Brief but efficient: acute HIV infection and the sexual transmission of HIV. *J Infect Dis* 2004,**189**:1785-1792.
289. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika C*, et al.* New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. *Lancet* 2008,**371**:2183-2191.
290. Buckner M. Manjako. In: *Encyclopedia of Sex and Gender: Men and Women in the World's Cultures*. Edited by Ember C, Ember M. New York: Kluwer Academic/Plenum; 2004. pp. 617-624.
291. Esbjornsson J, Mansson F, Martinez-Arias W, Vincic E, Biague AJ, da Silva ZJ*, et al.* Frequent CXCR4 tropism of HIV-1 subtype A and CRF02_AG during late-stage disease--indication of an evolving epidemic in West Africa. *Retrovirology* 2010,**7**:23.
292. Kong R, Li, H., Decker, J., Bibollet-Ruche, F., Kiviat, N., Hahn, B. and Shaw, G. Epitope Mapping of Neutralizing Antibody 1.7A on the HIV-2 gp160 Glycoprotein. In: *16th Conference on Retroviruses and Opportunistic Infections*. Montreal; 2009.
293. Posada D. ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. *Nucleic Acids Res* 2006,**34**:W700-703.
294. Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C*, et al.* Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* 2004,**14**:1229-1246.
295. Korber BT, Kunstman KJ, Patterson BK, Furtado M, McEvilly MM, Levy R*, et al.* Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J Virol* 1994,**68**:7467-7481.
296. Pond SL, Frost SD. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 2005,**21**:2531-2533.
297. Pond SL, Frost SD, Grossman Z, Gravenor MB, Richman DD, Brown AJ. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Comput Biol* 2006,**2**:e62.
298. Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 2005,**22**:1208-1222.
299. Delport W, Scheffler K, Botha G, Gravenor MB, Muse SV, Kosakovsky Pond SL. CodonTest: modeling amino acid substitution preferences in coding sequences. *PLoS Comput Biol* 2010,**6**.
300. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 2006,**23**:1891-1901.

301. Louder MK, Sambor A, Chertova E, Hunte T, Barrett S, Ojong F*, et al.* HIV-1 envelope pseudotyped viral vectors and infectious molecular clones expressing the same envelope glycoprotein have a similar neutralization phenotype, but culture in peripheral blood mononuclear cells is associated with decreased neutralization sensitivity. *Virology* 2005,**339**:226-238.
302. Rong R, Li B, Lynch RM, Haaland RE, Murphy MK, Mulenga J*, et al.* Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. *PLoS Pathog* 2009,**5**:e1000594.
303. Marcelino JM, Borrego P, Rocha C, Barroso H, Quintas A, Novo C*, et al.* Potent and broadly reactive HIV-2 neutralizing antibodies elicited by a vaccinia virus vector prime-C2V3C3 polypeptide boost immunization strategy. *J Virol* 2010,**84**:12429-12436.
304. Clapham PR, McKnight A, Weiss RA. Human immunodeficiency virus type 2 infection and fusion of CD4-negative human cell lines: induction and enhancement by soluble CD4. *J Virol* 1992,**66**:3531-3537.
305. Aasa-Chapman MM, Seymour CR, Williams I, McKnight A. Novel envelope determinants for CCR3 use by human immunodeficiency virus. *J Virol* 2006,**80**:10884-10889.
306. Frost SD, Wrin T, Smith DM, Kosakovsky Pond SL, Liu Y, Paxinos E*, et al.* Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection. *Proc Natl Acad Sci U S A* 2005,**102**:18514-18519.
307. Barroso H, Taveira N. Evidence for negative selective pressure in HIV-2 evolution in vivo. *Infect Genet Evol* 2005,**5**:239-246.
308. Deeks SG, Schweighardt B, Wrin T, Galovich J, Hoh R, Sinclair E*, et al.* Neutralizing antibody responses against autologous and heterologous viruses in acute versus chronic human immunodeficiency virus (HIV) infection: evidence for a constraint on the ability of HIV to completely evade neutralizing antibody responses. *J Virol* 2006,**80**:6155-6164.
309. Montefiori DC, Hill TS, Vo HT, Walker BD, Rosenberg ES. Neutralizing antibodies associated with viremia control in a subset of individuals after treatment of acute human immunodeficiency virus type 1 infection. *J Virol* 2001,**75**:10200-10207.
310. Kong R, Li H, Bibollet-Ruche F, Decker JM, Zheng NN, Gottlieb GS*, et al.* Broad and potent neutralizing antibody responses elicited in natural HIV-2 infection. *J Virol* 2012,**86**:947-960.
311. Bunnik EM, Pisas L, van Nuenen AC, Schuitemaker H. Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection. *J Virol* 2008,**82**:7932-7941.
312. Seaman MS, Janes H, Hawkins N, Grandpre LE, Devoy C, Giri A*, et al.* Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J Virol* 2010,**84**:1439-1452.
313. Wu X, Parast AB, Richardson BA, Nduati R, John-Stewart G, Mbori-Ngacha D*, et al.* Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant. *J Virol* 2006,**80**:835-844.
314. Skrabal K, Saragosti S, Labernardiere JL, Barin F, Clavel F, Mammano F. Human immunodeficiency virus type 1 variants isolated from single plasma samples display a wide spectrum of neutralization sensitivity. *J Virol* 2005,**79**:11848-11857.

315. Aasa-Chapman MM, Cheney K, Hue S, Forsman A, O'Farrell S, Pellegrino P*, et al. In Vivo* Emergence of HIV-1 Highly Sensitive to Neutralizing Antibodies. *PLoS ONE* 2011,**6(8):e23961**.
316. Moore PL, Ranchobe N, Lambson BE, Gray ES, Cave E, Abrahams MR*, et al.* Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection. *PLoS Pathog* 2009,**5**:e1000598.
317. Davis KL, Bibollet-Ruche F, Li H, Decker JM, Kutsch O, Morris L*, et al.* Human immunodeficiency virus type 2 (HIV-2)/HIV-1 envelope chimeras detect high titers of broadly reactive HIV-1 V3-specific antibodies in human plasma. *J Virol* 2009,**83**:1240-1259.
318. Calado M, Matoso P, Santos-Costa Q, Espirito-Santo M, Machado J, Rosado L*, et al.* Coreceptor usage by HIV-1 and HIV-2 primary isolates: the relevance of CCR8 chemokine receptor as an alternative coreceptor. *Virology* 2010,**408**:174-182.
319. Sagar M, Wu X, Lee S, Overbaugh J. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J Virol* 2006,**80**:9586-9598.
320. Moore PL, Gray ES, Morris L. Specificity of the autologous neutralizing antibody response. *Curr Opin HIV AIDS* 2009,**4**:358-363.
321. Rong R, Bibollet-Ruche F, Mulenga J, Allen S, Blackwell JL, Derdeyn CA. Role of V1V2 and other human immunodeficiency virus type 1 envelope domains in resistance to autologous neutralization during clade C infection. *J Virol* 2007,**81**:1350-1359.
322. Pinter A, Honnen WJ, He Y, Gorny MK, Zolla-Pazner S, Kayman SC. The V1/V2 domain of gp120 is a global regulator of the sensitivity of primary human immunodeficiency virus type 1 isolates to neutralization by antibodies commonly induced upon infection. *J Virol* 2004,**78**:5205-5215.
323. Meyerhans A, Vartanian JP, Wain-Hobson S. DNA recombination during PCR. *Nucleic Acids Res* 1990,**18**:1687-1691.
324. Jordan MR, Kearney M, Palmer S, Shao W, Maldarelli F, Coakley EP*, et al.* Comparison of standard PCR/cloning to single genome sequencing for analysis of HIV-1 populations. *J Virol Methods* 2010,**168**:114-120.
325. Fenyo EM, Heath A, Dispinseri S, Holmes H, Lusso P, Zolla-Pazner S*, et al.* International network for comparison of HIV neutralization assays: the NeutNet report. *PLoS One* 2009,**4**:e4505.
326. Martin N, Welsch S, Jolly C, Briggs JA, Vaux D, Sattentau QJ. Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. *J Virol* 2010,**84**:3516-3527.
327. Carlson JM, Brumme ZL. HIV evolution in response to HLA-restricted CTL selection pressures: a population-based perspective. *Microbes Infect* 2008,**10**:455-461.
328. McMichael AJ, Phillips RE. Escape of human immunodeficiency virus from immune control. *Annu Rev Immunol* 1997,**15**:271-296.
329. Blish CA, Blay WM, Haigwood NL, Overbaugh J. Transmission of HIV-1 in the face of neutralizing antibodies. *Curr HIV Res* 2007,**5**:578-587.
330. van der Kuyl AC, Cornelissen M. Identifying HIV-1 dual infections. *Retrovirology* 2007,**4**:67.
331. Aasa-Chapman MM, Holuigue S, Aubin K, Wong M, Jones NA, Cornforth D*, et al.* Detection of antibody-dependent complement-mediated inactivation of

both autologous and heterologous virus in primary human immunodeficiency virus type 1 infection. *J Virol* 2005,**79**:2823-2830.
332. Wren L, Kent SJ. HIV Vaccine efficacy trial: Glimmers of hope and the potential role of antibody-dependent cellular cytotoxicity. *Hum Vaccin* 2011,**7**.
333. Chahroudi A, Bosinger SE, Vanderford TH, Paiardini M, Silvestri G. Natural SIV hosts: showing AIDS the door. *Science* 2012,**335**:1188-1193.
334. de Silva TI, Cotten M, Rowland-Jones SL. HIV-2: the forgotten AIDS virus. *Trends Microbiol* 2008,**16**:588-595.
335. Hodges-Mameletzis I, De Bree GJ, Rowland-Jones SL. An underestimated lentivirus model: what can HIV-2 research contribute to the development of an effective HIV-1 vaccine? *Expert Rev Anti Infect Ther* 2011,**9**:195-206.
336. Tienen C, van der Loeff MS, Zaman SM, Vincent T, Sarge-Njie R, Peterson I*, et al.* Two distinct epidemics: the rise of HIV-1 and decline of HIV-2 infection between 1990 and 2007 in rural Guinea-Bissau. *J Acquir Immune Defic Syndr* 2010,**53**:640-647.
337. Holmgren B, Aaby P, Jensen H, Larsen O, da Silva Z, Lisse IM. Increased prevalence of retrovirus infections among older women in Africa. *Scand J Infect Dis* 1999,**31**:459-466.
338. Brinkhof MW, Boulle A, Weigel R, Messou E, Mathers C, Orrell C*, et al.* Mortality of HIV-infected patients starting antiretroviral therapy in sub-Saharan Africa: comparison with HIV-unrelated mortality. *PLoS Med* 2009,**6**:e1000066.
339. Dong T, Stewart-Jones G, Chen N, Easterbrook P, Xu X, Papagno L*, et al.* HIV-specific cytotoxic T cells from long-term survivors select a unique T cell receptor. *J Exp Med* 2004,**200**:1547-1557.
340. Appay V, Nixon DF, Donahoe SM, Gillespie GM, Dong T, King A*, et al.* HIV-specific CD8(+) T cells produce antiviral cytokines but are impaired in cytolytic function. *J Exp Med* 2000,**192**:63-75.
341. Kostense S, Ogg GS, Manting EH, Gillespie G, Joling J, Vandenberghe K*, et al.* High viral burden in the presence of major HIV-specific CD8(+) T cell expansions: evidence for impaired CTL effector function. *Eur J Immunol* 2001,**31**:677-686.
342. Appay V, Dunbar PR, Callan M, Klenerman P, Gillespie GM, Papagno L*, et al.* Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. *Nat Med* 2002,**8**:379-385.
343. Tussey LG, Nair US, Bachinsky M, Edwards BH, Bakari J, Grimm K*, et al.* Antigen burden is major determinant of human immunodeficiency virus-specific CD8+ T cell maturation state: potential implications for therapeutic immunization. *J Infect Dis* 2003,**187**:364-374.
344. Yamamoto T, Price DA, Casazza JP, Ferrari G, Nason M, Chattopadhyay PK*, et al.* Surface expression patterns of negative regulatory molecules identify determinants of virus-specific CD8+ T-cell exhaustion in HIV infection. *Blood* 2011,**117**:4805-4815.
345. Hersperger AR, Martin JN, Shin LY, Sheth PM, Kovacs CM, Cosma GL*, et al.* Increased HIV-specific CD8+ T-cell cytotoxic potential in HIV elite controllers is associated with T-bet expression. *Blood* 2011,**117**:3799-3808.
346. Liu Z, Cumberland WG, Hultin LE, Prince HE, Detels R, Giorgi JV. Elevated CD38 antigen expression on CD8+ T cells is a stronger marker for the risk of chronic HIV disease progression to AIDS and death in the Multicenter AIDS Cohort Study than CD4+ cell count, soluble immune activation markers, or

combinations of HLA-DR and CD38 expression. *J Acquir Immune Defic Syndr Hum Retrovirol* 1997,**16**:83-92.
347. Chun TW, Justement JS, Sanford C, Hallahan CW, Planta MA, Loutfy M*, et al.* Relationship between the frequency of HIV-specific CD8+ T cells and the level of CD38+CD8+ T cells in untreated HIV-infected individuals. *Proc Natl Acad Sci U S A* 2004,**101**:2464-2469.
348. Lopez M, Soriano V, Benito JM. Escape mutations in HIV infection and its impact on CD8+ T cell responses. *Curr Mol Med* 2007,**7**:446-458.
349. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E*, et al.* CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med* 2007,**13**:46-53.
350. Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D*, et al.* Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* 2008,**372**:1881-1893.
351. Chevalier MF, Julg B, Pyo A, Flanders M, Ranasinghe S, Soghoian DZ*, et al.* HIV-1-specific interleukin-21+ CD4+ T cell responses contribute to durable viral control through the modulation of HIV-specific CD8+ T cell function. *J Virol* 2011,**85**:733-741.
352. Zuniga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, Hernandez A*, et al.* Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. *J Virol* 2006,**80**:3122-3125.
353. Masemola A, Mashishi T, Khoury G, Mohube P, Mokgotho P, Vardas E*, et al.* Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. *J Virol* 2004,**78**:3233-3243.
354. Navis M, Schellens I, van Baarle D, Borghans J, van Swieten P, Miedema F*, et al.* Viral replication capacity as a correlate of HLA B57/B5801-associated nonprogressive HIV-1 infection. *J Immunol* 2007,**179**:3133-3143.
355. Julg B, Williams KL, Reddy S, Bishop K, Qi Y, Carrington M*, et al.* Enhanced anti-HIV functional activity associated with Gag-specific CD8 T-cell responses. *J Virol* 2010,**84**:5540-5549.
356. Kostense S, Vandenberghe K, Joling J, Van Baarle D, Nanlohy N, Manting E*, et al.* Persistent numbers of tetramer+ CD8(+) T cells, but loss of interferon-gamma+ HIV-specific T cells during progression to AIDS. *Blood* 2002,**99**:2505-2511.
357. Holtappels R, Janda J, Thomas D, Schenk S, Reddehase MJ, Geginat G. Adoptive CD8 T cell control of pathogens cannot be improved by combining protective epitope specificities. *J Infect Dis* 2008,**197**:622-629.
358. Addo MM, Draenert R, Rathod A, Verrill CL, Davis BT, Gandhi RT*, et al.* Fully differentiated HIV-1 specific CD8+ T effector cells are more frequently detectable in controlled than in progressive HIV-1 infection. *PLoS One* 2007,**2**:e321.
359. Papagno L, Appay V, Sutton J, Rostron T, Gillespie GM, Ogg GS*, et al.* Comparison between HIV- and CMV-specific T cell responses in long-term HIV infected donors. *Clin Exp Immunol* 2002,**130**:509-517.
360. Burgers WA, Riou C, Mlotshwa M, Maenetje P, de Assis Rosa D, Brenchley J*, et al.* Association of HIV-specific and total CD8+ T memory phenotypes in subtype C HIV-1 infection with viral set point. *J Immunol* 2009,**182**:4751-4761.

361. Sauce D, Almeida JR, Larsen M, Haro L, Autran B, Freeman GJ*, et al.* PD-1 expression on human CD8 T cells depends on both state of differentiation and activation status. *AIDS* 2007,**21**:2005-2013.
362. Ozkaya Sahin G, Holmgren B, da Silva Z, Nielsen J, Nowroozalizadeh S, Esbjornsson J*, et al.* Potent intratype neutralizing activity distinguishes human immunodeficiency virus type 2 (HIV-2) from HIV-1. *J Virol* 2012,**86**:961-971.
363. Barouch DH, Liu J, Li H, Maxfield LF, Abbink P, Lynch DM*, et al.* Vaccine protection against acquisition of neutralization-resistant SIV challenges in rhesus monkeys. *Nature* 2012,**482**:89-93.
364. Keele BF, Jones JH, Terio KA, Estes JD, Rudicell RS, Wilson ML*, et al.* Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* 2009,**460**:515-519.

APPENDIX

Appendix 1. Table of all sequence details used in Chapter 3

Sample ID	Country[1]	Year of sampling	Virus Region	Accession Number
CA5625	GWc	2007	HIV-2 p26	submitted
CA6417	GWc	2007	HIV-2 p26	submitted
CA6666	GWc	2007	HIV-2 p26	submitted
CA6936	GWc	2007	HIV-2 p26	submitted
CA7164	GWc	2007	HIV-2 p26	submitted
CA7205	GWc	2007	HIV-2 p26	submitted
CA7235	GWc	2007	HIV-2 p26	submitted
CA7253	GWc	2007	HIV-2 p26	submitted
CA7284	GWc	2007	HIV-2 p26	submitted
CA7340	GWc	2007	HIV-2 p26	submitted
CA8036	GWc	2008	HIV-2 p26	submitted
N00009	GWc	1996	HIV-2 p26	submitted
N00038	GWc	1996	HIV-2 p26	submitted
N00054	GWc	1996	HIV-2 p26	submitted
N00076	GWc	1996	HIV-2 p26	submitted
N00096	GWc	1996	HIV-2 p26	submitted
N00115	GWc	1996	HIV-2 p26	submitted
N00125	GWc	1996	HIV-2 p26	submitted
N00126	GWc	1996	HIV-2 p26	submitted
N00129	GWc	1996	HIV-2 p26	submitted
N00135	GWc	1996	HIV-2 p26	submitted
N00154	GWc	1996	HIV-2 p26	submitted
N00157	GWc	1996	HIV-2 p26	submitted
N00186	GWc	1996	HIV-2 p26	submitted
N00203	GWc	1996	HIV-2 p26	submitted
N00255	GWc	1996	HIV-2 p26	submitted
N04066	GWc	2003	HIV-2 p26	submitted
N04083	GWc	2003	HIV-2 p26	submitted
N04189	GWc	2003	HIV-2 p26	submitted
N04265	GWc	2003	HIV-2 p26	submitted
N04269	GWc	2003	HIV-2 p26	submitted
N04276	GWc	2003	HIV-2 p26	submitted
N04280	GWc	2003	HIV-2 p26	submitted
N04388	GWc	2003	HIV-2 p26	submitted
N04405	GWc	2003	HIV-2 p26	submitted
N04426	GWc	2003	HIV-2 p26	submitted
N04722	GWc	2008	HIV-2 p26	submitted
N62522	GWc	1996	HIV-2 p26	submitted

N65309	GWc	2006	HIV-2 p26	GQ485448
N65310	GWc	2006	HIV-2 p26	GQ485449
N65313	GWc	2006	HIV-2 p26	GQ485450
N65330	GWc	2006	HIV-2 p26	GQ485457
N65331	GWc	2006	HIV-2 p26	GQ485458
N65333	GWc	2006	HIV-2 p26	GQ485460
N65336	GWc	2006	HIV-2 p26	GQ485461
N65341	GWc	2006	HIV-2 p26	GQ485465
N65347	GWc	2006	HIV-2 p26	GQ485467
N65349	GWc	2006	HIV-2 p26	GQ485468
N65350	GWc	2006	HIV-2 p26	GQ485469
N65354	GWc	2006	HIV-2 p26	GQ485472
N65358	GWc	2006	HIV-2 p26	GQ485475
N65361	GWc	2006	HIV-2 p26	GQ485476
N65363	GWc	2006	HIV-2 p26	GQ485477
N65367	GWc	2006	HIV-2 p26	GQ485478
N65369	GWc	2006	HIV-2 p26	GQ485479
N65370	GWc	2006	HIV-2 p26	GQ485480
N65382	GWc	2006	HIV-2 p26	GQ485482
N65384	GWc	2006	HIV-2 p26	GQ485483
N65386	GWc	2006	HIV-2 p26	GQ485484
N65388	GWc	2006	HIV-2 p26	GQ485485
N65392	GWc	2006	HIV-2 p26	GQ485488
N65394	GWc	2006	HIV-2 p26	GQ485489
N65398	GWc	2006	HIV-2 p26	GQ485490
N65403	GWc	2006	HIV-2 p26	GQ485491
N65406	GWc	2006	HIV-2 p26	GQ485492
N65410	GWc	2006	HIV-2 p26	GQ485495
N65415	GWc	2006	HIV-2 p26	GQ485496
N65418	GWc	2006	HIV-2 p26	GQ485497
N65424	GWc	2006	HIV-2 p26	GQ485498
N65426	GWc	2006	HIV-2 p26	GQ485499
N65435	GWc	2006	HIV-2 p26	GQ485500
N65436	GWc	2006	HIV-2 p26	GQ485501
N65494	GWc	2006	HIV-2 p26	GQ485504
N65500	GWc	2006	HIV-2 p26	GQ485505
N65552	GWc	2006	HIV-2 p26	GQ485506
N65555	GWc	2006	HIV-2 p26	GQ485507
N65575	GWc	2006	HIV-2 p26	GQ485509
N65600	GWc	2006	HIV-2 p26	submitted
N65601	GWc	2006	HIV-2 p26	GQ485511
N65605	GWc	2006	HIV-2 p26	GQ485512
N65611	GWc	2006	HIV-2 p26	GQ485513
N65612	GWc	2006	HIV-2 p26	GQ485514
N65613	GWc	2006	HIV-2 p26	GQ485515
N65666	GWc	2007	HIV-2 p26	submitted

N65667	GWc	2007	HIV-2 p26	submitted
N65668	GWc	2007	HIV-2 p26	submitted
9649	ML	1995	HIV-2 env	AF170039
96150	CV	1996	HIV-2 env	AF170041
96151	ML	1996	HIV-2 env	AF170034
96199	CV	1995	HIV-2 env	AF170043
96202	BF	1995	HIV-2 env	AF170040
96203	CV	1995	HIV-2 env	AF170049
96205	GH	1996	HIV-2 env	AF170031
96206	SN	1994	HIV-2 env	AF170048
96308	CV	1994	HIV-2 env	AF170042
96310	SN	1996	HIV-2 env	AF170036
96323	CI	1994	HIV-2 env	AF170032
96324	GW	1996	HIV-2 env	AF170046
96326	ML	1994	HIV-2 env	AF170038
96327	CV	1996	HIV-2 env	AF170044
96329	GW	1994	HIV-2 env	AF170045
97223	CI	1997	HIV-2 env	AF170033
03CM510	CM	2003	HIV-2 env	EU028345
7312A	CI	1990	HIV-2 env	L36874
ALI	GW	1989	HIV-2 env	AF082339
B1002	GWc	1991	HIV-2 env	AJ008283
B1010	GWc	1991	HIV-2 env	AJ008284
B1011	GWc	1991	HIV-2 env	AJ011223
B1014	GWc	1991	HIV-2 env	AJ011250
B1015	GWc	1991	HIV-2 env	AJ008285
B1022	GWc	1991	HIV-2 env	AJ008286
B1024	GWc	1991	HIV-2 env	AJ011224
B1025	GWc	1991	HIV-2 env	AJ011225
B1027	GWc	1991	HIV-2 env	AJ011226
B1028	GWc	1991	HIV-2 env	AJ008288
B1033	GWc	1991	HIV-2 env	AJ008289
B1036	GWc	1991	HIV-2 env	AJ011227
B1041	GWc	1991	HIV-2 env	AJ008290
B1042	GWc	1991	HIV-2 env	AJ011228
B1046	GWc	1991	HIV-2 env	AJ008291
B1048	GWc	1991	HIV-2 env	AJ008292
B1049	GWc	1991	HIV-2 env	AJ008293
B1053	GWc	1991	HIV-2 env	AJ011251
B1054	GWc	1991	HIV-2 env	AJ011231
B1058	GWc	1991	HIV-2 env	AJ008294
B1068	GWc	1991	HIV-2 env	AJ008295
B1079	GWc	1991	HIV-2 env	AJ008297
B1083	GWc	1991	HIV-2 env	AJ011252
B1084	GWc	1991	HIV-2 env	AJ011253
B1088	GWc	1991	HIV-2 env	AJ011254

B1089	GWc	1991	HIV-2 env	AJ011255
B1095	GWc	1991	HIV-2 env	AJ008298
B1100	GWc	1991	HIV-2 env	AJ011233
B1121	GWc	1991	HIV-2 env	AJ008299
B1123	GWc	1991	HIV-2 env	AJ011256
B1124	GWc	1991	HIV-2 env	AJ011257
B1127	GWc	1991	HIV-2 env	AJ011234
B1128	GWc	1991	HIV-2 env	AJ011235
B1129	GWc	1991	HIV-2 env	AJ008300
B1133	GWc	1991	HIV-2 env	AJ011258
B1134	GWc	1991	HIV-2 env	AJ011237
B1146	GWc	1991	HIV-2 env	AJ008301
B1153	GWc	1991	HIV-2 env	AJ011238
B1160	GWc	1991	HIV-2 env	AJ011239
B1165	GWc	1991	HIV-2 env	AJ008302
B1174	GWc	1991	HIV-2 env	AJ011260
B1183	GWc	1991	HIV-2 env	AJ008303
B1187	GWc	1991	HIV-2 env	AJ011262
B1192	GWc	1991	HIV-2 env	AJ008304
B1193	GWc	1991	HIV-2 env	AJ011242
B1198	GWc	1991	HIV-2 env	AJ011264
B1200	GWc	1991	HIV-2 env	AJ011265
B1207	GWc	1991	HIV-2 env	AJ011243
B1210	GWc	1991	HIV-2 env	AJ011244
B1215	GWc	1991	HIV-2 env	AJ008305
B1216	GWc	1991	HIV-2 env	AJ011245
B1222	GWc	1991	HIV-2 env	AJ011267
B1229	GWc	1991	HIV-2 env	AJ008306
B1231	GWc	1991	HIV-2 env	AJ008307
B1233	GWc	1991	HIV-2 env	AJ011268
B1242	GWc	1991	HIV-2 env	AJ008309
B1255	GWc	1991	HIV-2 env	AJ008310
B1262	GWc	1991	HIV-2 env	AJ011248
B1265	GWc	1991	HIV-2 env	AJ008311
B1270	GWc	1991	HIV-2 env	AJ011269
B1276	GWc	1991	HIV-2 env	AJ011270
B1277	GWc	1991	HIV-2 env	AJ008312
B1286	GWc	1991	HIV-2 env	AJ011271
B1293	GWc	1991	HIV-2 env	AJ008313
B1297	GWc	1991	HIV-2 env	AJ008314
B1301	GWc	1991	HIV-2 env	AJ011272
B1302	GWc	1991	HIV-2 env	AJ008315
B1303	GWc	1991	HIV-2 env	AJ008316
BEN	ML	1987	HIV-2 env	M30502
CA6417	GWc	2007	HIV-2 env	JN863883
CA7164	GWc	2007	HIV-2 env	JN863884

CA7205	GWc	2007	HIV-2 env	JN863897
CA7253	GWc	2007	HIV-2 env	JN863898
CA7340	GWc	2007	HIV-2 env	JN863882
CAM1	GW	1987	HIV-2 env	U05359
CAM2	GW	1987	HIV-2 env	D00835
CAM3	GW	1987	HIV-2 env	U05355
CAM4	GW	1987	HIV-2 env	U05356
CAM5	GW	1987	HIV-2 env	U05357
CAM6	GW	1987	HIV-2 env	U05358
CBL21	GM	1988	HIV-2 env	U05350
CBL22	GM	1988	HIV-2 env	U05351
CBL23	GM	1988	HIV-2 env	AY509259
CBL24	GM	1988	HIV-2 env	AJ238999
D194	GM	1987	HIV-2 env	J04542
FGNIHZ	GW	1986	HIV-2 env	J03654
GH1	CI	1986	HIV-2 env	M30895
ISYSBL6669	GM	1985	HIV-2 env	J04498
MIC97	GW	1997	HIV-2 env	AY168925
MJC97	GW	1997	HIV-2 env	EU021092
N4013	GWc	2003	HIV-2 env	JN863846
N4102	GWc	2003	HIV-2 env	JN863856
N4115	GWc	2003	HIV-2 env	JN863860
N4137	GWc	2003	HIV-2 env	JN863861
N4145	GWc	2003	HIV-2 env	JN863854
N4153	GWc	2003	HIV-2 env	JN863863
N4167	GWc	2003	HIV-2 env	JN863862
N4172	GWc	2003	HIV-2 env	JN863867
N4182	GWc	2003	HIV-2 env	JN863864
N4205	GWc	2003	HIV-2 env	JN863858
N4206	GWc	2003	HIV-2 env	JN863865
N4218	GWc	2003	HIV-2 env	JN863857
N4220	GWc	2003	HIV-2 env	JN863855
N4258	GWc	2003	HIV-2 env	JN863852
N4269	GWc	2003	HIV-2 env	JN863836
N4272	GWc	2003	HIV-2 env	JN863810
N4289	GWc	2003	HIV-2 env	JN863792
N4291	GWc	2003	HIV-2 env	JN863853
N4292	GWc	2003	HIV-2 env	JN863859
N4388	GWc	2003	HIV-2 env	JN863866
N4401	GWc	2003	HIV-2 env	JN863801
N4722	GWc	2008	HIV-2 env	JN863885
N65317	GWc	2006	HIV-2 env	JN863868
N65319	GWc	2006	HIV-2 env	GQ485518
N65320	GWc	2006	HIV-2 env	JN863869
N65326	GWc	2006	HIV-2 env	GQ485519
N65330	GWc	2006	HIV-2 env	GQ485520

N65332	GWc	2006	HIV-2 env	GQ485522
N65333	GWc	2006	HIV-2 env	GQ485523
N65338	GWc	2006	HIV-2 env	GQ485524
N65347	GWc	2006	HIV-2 env	GQ485526
N65349	GWc	2006	HIV-2 env	GQ485527
N65354	GWc	2006	HIV-2 env	GQ485528
N65355	GWc	2006	HIV-2 env	GQ485529
N65358	GWc	2006	HIV-2 env	GQ485530
N65361	GWc	2006	HIV-2 env	GQ485531
N65363	GWc	2006	HIV-2 env	GQ485532
N65365	GWc	2006	HIV-2 env	JN863870
N65367	GWc	2006	HIV-2 env	GQ485533
N65369	GWc	2006	HIV-2 env	GQ485534
N65370	GWc	2006	HIV-2 env	GQ485535
N65371	GWc	2006	HIV-2 env	JN863871
N65372	GWc	2006	HIV-2 env	JN863872
N65373	GWc	2006	HIV-2 env	GQ485536
N65381	GWc	2006	HIV-2 env	JN863873
N65382	GWc	2006	HIV-2 env	GQ485537
N65388	GWc	2006	HIV-2 env	JN863874
N65389	GWc	2006	HIV-2 env	GQ485538
N65390	GWc	2006	HIV-2 env	JN863875
N65391	GWc	2006	HIV-2 env	GQ485539
N65392	GWc	2006	HIV-2 env	GQ485540
N65395	GWc	2006	HIV-2 env	JN863876
N65399	GWc	2006	HIV-2 env	JN863877
N65402	GWc	2006	HIV-2 env	JN863878
N65409	GWc	2006	HIV-2 env	GQ485543
N65411	GWc	2006	HIV-2 env	JN863879
N65415	GWc	2006	HIV-2 env	GQ485544
N65418	GWc	2006	HIV-2 env	GQ485545
N65426	GWc	2006	HIV-2 env	GQ485546
N65494	GWc	2006	HIV-2 env	GQ485547
N65552	GWc	2006	HIV-2 env	JN863880
N65601	GWc	2006	HIV-2 env	GQ485548
N65605	GWc	2006	HIV-2 env	GQ485549
N65613	GWc	2006	HIV-2 env	GQ485550
N65614	GWc	2006	HIV-2 env	JN863881
PEI2	ML	1987	HIV-2 env	U22047
ROD	CV	1985	HIV-2 env	M15390
ST	SN	1987	HIV-2 env	M31113
UC2	CI	1988	HIV-2 env	U38293
vcp	GW	1986	HIV-2 env	EU580099
E134	GWc	1996	HIV-1 env	JN863732
E159	GWc	1996	HIV-1 env	JN863733
E197	GWc	1996	HIV-1 env	JN863734

E246	GWc	1996	HIV-1 env	JN863735
E4005	GWc	2003	HIV-1 env	JN863773
E4006	GWc	2003	HIV-1 env	JN863772
E4015	GWc	2003	HIV-1 env	JN863774
E4037	GWc	2003	HIV-1 env	JN863775
E4048	GWc	2003	HIV-1 env	JN863776
E4078	GWc	2003	HIV-1 env	JN863777
E4080	GWc	2003	HIV-1 env	JN863778
E4084	GWc	2003	HIV-1 env	JN863771
E4101	GWc	2003	HIV-1 env	JN863779
E4106	GWc	2003	HIV-1 env	JN863781
E4107	GWc	2003	HIV-1 env	JN863780
E4128	GWc	2003	HIV-1 env	JN863782
E4129	GWc	2003	HIV-1 env	JN863770
E4163	GWc	2003	HIV-1 env	JN863769
E4171	GWc	2003	HIV-1 env	JN863783
E4174	GWc	2003	HIV-1 env	JN863784
E4197	GWc	2003	HIV-1 env	JN863768
E4212	GWc	2003	HIV-1 env	JN863785
E4233	GWc	2003	HIV-1 env	JN863767
E4285	GWc	2003	HIV-1 env	JN863786
E4301	GWc	2003	HIV-1 env	JN863766
E4366	GWc	2003	HIV-1 env	JN863765
E4385	GWc	2003	HIV-1 env	JN863736
E4394	GWc	2003	HIV-1 env	JN863787
E4418	GWc	2003	HIV-1 env	JN863764
E4424	GWc	2003	HIV-1 env	JN863763
E4425	GWc	2003	HIV-1 env	JN863762
E5006	GWc	2006	HIV-1 env	JN863737
E5050	GWc	2006	HIV-1 env	JN863738
E5380	GWc	2006	HIV-1 env	JN863739
E5535	GWc	2006	HIV-1 env	JN863740
E5539	GWc	2006	HIV-1 env	JN863741
E5657	GWc	2006	HIV-1 env	JN863742
E6006	GWc	2006	HIV-1 env	JN863743
E6403	GWc	2006	HIV-1 env	JN863744
E6482	GWc	2007	HIV-1 env	JN863745
E6498	GWc	2007	HIV-1 env	JN863746
E6591	GWc	2007	HIV-1 env	JN863747
E6865	GWc	2007	HIV-1 env	JN863748
E6922	GWc	2007	HIV-1 env	JN863749
E6952	GWc	2007	HIV-1 env	JN863750
E6958	GWc	2007	HIV-1 env	JN863751
E7019	GWc	2006	HIV-1 env	JN863752
E7114	GWc	2006	HIV-1 env	JN863753
E7165	GWc	2006	HIV-1 env	JN863754

E7367	GWc	2007	HIV-1 env	JN863755
E7424	GWc	2007	HIV-1 env	JN863756
E7467	GWc	2007	HIV-1 env	JN863757
E7582	GWc	2007	HIV-1 env	JN863758
E7635	GWc	2007	HIV-1 env	JN863759
E7650	GWc	2007	HIV-1 env	JN863760
E7743	GWc	2007	HIV-1 env	JN863761

[1] County code: BF: Burkina Faso, CI: Cote d'Ivoire, GH: Ghana, GM: Gambia, CV: Cape Verde, GW: Guinea Bissau, GWc: Guinea Bissau, Caió., ML: Mali, SN: Senegal, CM: Cameroon.

Appendix 2. HLA B allele distribution in HIV-2 infected subjects (where available, N = 56). Subjects in bold denote those with appropriate HLA B alleles (B14, B35, B53, B58) in whom tetramer staining was attempted. ^ Individuals with insufficient sample volume left for tetramer experiments.

Subject	CD4 count (cells/µl)	Viral load (copies/ml)	HLA B1	HLA B2
TD02	**525**	**<100**	**B*3501**	**B*5301**
TD03	560	82005	B*1302	B*5101
TD05	179	393993	B*0801	B*1510
TD06	**1176**	**<100**	**B*5301**	**B*5702**
TD07	**1135**	**<100**	**B*0801**	**B*5301**
TD08	**465**	**3433**	**B*5301**	**B*5703**
TD09^	314	693	B*4901	B*5301
TD11	709	<100	B*1503	B*1510
TD12	347	865	B*0801	B*3910
TD13	**509**	**1632**	**B*1402**	**B*1516**
TD15	**800**	**110**	**B*1402**	**B*5801**
TD16	**1313**	**<100**	**B*5301**	**B*5703**
TD17	572	<100	B*1510	B*1510
TD20	**823**	**<100**	**B*5301**	**B*5801**
TD21	952	793	B*1302	B*5001
TD22	444	<100	B*4901	B*8201
TD23	**369**	**517**	**B*1302**	**B*1402**
TD24	**191**	**10560**	**B*3501**	**B*4901**
TD25	**669**	**<100**	**B*1510**	**B*5301**
TD26	**356**	**892**	**B*1402**	**B*5101**
TD27	**435**	**<100**	**B*3501**	**B*4901**
TD28	**724**	**<100**	**B*1510**	**B*5301**
TD29	521	<100	B*1503	B*5703

TD30^	890	<100	B*1402	B*8201
TD31	407	107183	B*4701	B*5001
TD32	616	<100	B*1510	B*5201
TD34	530	734	B*1503	B*4103
TD35	**765**	**<100**	**B*0801**	**B*3501**
TD36^	489	<100	B*0801	B*5801
TD37	**293**	**708**	**B*0801**	**B*3501**
TD38	**977**	**<100**	**B*5801**	**B*5801**
TD39	**581**	**<100**	**B*3501**	**B*3501**
TD40	**673**	**<100**	**B*0801**	**B*5801**
TD41	213	20956	B*0801	B*1510
TD42	**675**	**200**	**B*0801**	**B*5801**
TD43	382	<100	B*1510	B*4403
TD44	**1063**	**<100**	**B*5301**	**B*5801**
TD45	**596**	**<100**	**B*4901**	**B*5301**
TD46	**301**	**<100**	**B*1503**	**B*3501**
TD47	187	1174	B*0801	B*1503
TD48	**576**	**<100**	**B*3501**	**B*4201**
TD49	**311**	**666**	**B*3501**	**B*3501**
TD50	520	<100	B*4403	B*4901
TD51	447	<100	B*1510	B*1510
TD52	321	<100	B*0801	B*5001
TD54	**925**	**<100**	**B*5001**	**B*5801**
TD55	407	484	B*0801	B*8201
TD56	**965**	**<100**	**B*1402**	**B*5301**
TD57	**1127**	**690**	**B*1510**	**B*3501**
TD58	**312**	**<100**	**B*0702**	**B*3501**
TD59	**483**	**1391**	**B*0801**	**B*3501**
TD60	770	3878	B*0702	B*1503
TD61	497	<100	B*1510	B*4901
TD62	497	139519	B*0801	B*1302
TD92	347	<100	B*1503	B*4901
TD93	**400**	**763**	**B*5801**	**B*5801**